Transmisión de datos

Problemas resueltos

Mónica Aguilar Igartua
Jordi Forné Muñoz
Jordi Mata Díaz
Francisco Rico Novella
Alfonso Rojas Espinosa
Miquel Soriano Ibáñez

UPCGRAU 1

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Transmisión de datos

Problemas resueltos

Mónica Aguilar Igartua
Jordi Forné Muñoz
Jordi Mata Díaz
Francisco Rico Novella
Alfonso Rojas Espinosa
Miquel Soriano Ibáñez

UPCGRAU 1

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Primera edición: noviembre de 2010
Reedición: febrero de 2025

Oficina de Publicacions Acadèmiques Digitals de la UPC
Edificio K2M, Planta S1, Despacho S103-S104
Jordi Girona 1-3, 08034 Barcelona
Tel.: 934 015 885
www.upc.edu/idp
E-mail: info.idp@upc.edu

ISBN: 978-84-7653-514-1
DOI: 10.5821/ebook-9788476535141

Índice

Índice de figuras

1
Codificación de fuente

1.1. Introducción

La transmisión de datos es el conjunto de técnicas y conceptos que surgen al estudiar el problema de la transmisión de información digital, cualesquiera que sean su origen y naturaleza. La transmisión se realizará a través de un canal físico limitado en ancho de banda y potencia, como puede ser un par de cables, un cable coaxial, una fibra óptica, un radioenlace, o una combinación de estos.

Una descripción global de lo que constituye la transmisión de datos debe comenzar con la distinción conceptual de los diferentes elementos de que se compone. Esta división permitirá una mayor comprensión del problema y, consecuentemente, una mayor capacidad de análisis.

El primer paso es la compresión de las fuentes de datos (voz, imágenes, datos digitales, etc.) a partir de la definición del concepto de información realizada por Shannon [ABR63]. La formalización del concepto de información nos lleva, además, a estudiar el comportamiento de un sistema considerando la transmisión de secuencias de datos aleatorias. De esta forma, el problema inicial se ha dividido en dos: la caracterización de la fuente y la caracterización del canal, todo ello sin pérdida de generalidad.

En este capítulo, se trata el problema partiendo de una fuente discreta equivalente. En general, la transmisión de los datos tal como manan de la fuente conllevaría un derroche de recursos. Para reducir la redundancia, debemos recurrir a la compresión [HAN03]. Shannon establece un límite teórico por debajo del cual ya no puede comprimirse más sin pérdidas. Dicho límite depende de la estadística de emisión y se denomina entropía. La entropía es un parámetro básico y propio de la fuente.

Algunos codificadores de fuente requieren el conocimiento exacto y a priori de las características estadísticas de emisión, mientras que otros lo van adquiriendo de una

manera adaptativa a partir de los propios datos emitidos. Un ejemplo de los primeros es el codificador de Huffman, y de los segundos el codificador de Ziv-Lempel. En ambos casos después del proceso de compresión se obtiene una secuencia de bits independientes, que caracterizaremos mediante fuente binaria equivalente.

El siguiente proceso es el mapeo de estos bits en los símbolos que el alfabeto de entrada del sistema modulador, mediante la codificación elegida. En este punto, el problema se reduce a la transmisión de estos símbolos al receptor, que realizará el proceso de decodificación inverso convirtiéndolos en una secuencia de bits que idealmente coincidirá con la emitida. La máxima velocidad a la que esta secuencia de bits puede ser transmitida de forma fiable se denomina *capacidad del canal*, y fue también establecida por Shannon [COV06].

1.2. Contenidos teóricos

- Teoría de la información
  - Concepto de información
  - Entropía. Entropía conjunta. Entropía condicional
  - Información mutua
  - Entropía de una fuente con memoria
- Codificación
  - Códigos instantáneos
  - Códigos de Huffman
  - Códigos de ráfagas
  - Códigos aritméticos
  - Códigos diccionario
- Capacidad de canal
  - Caracterización de un canal discreto
  - Capacidad de un canal simétrico sin memoria

1.3. Bibliografía

[ABR63] Abranson, N., *Information Theory and Coding*, McGraw-Hill Education, ISBN-10: 0070001456, 1963

[HAN03] Hankerson, D.C.; Harris, G.; Johnson P.D., *Introduction to Information Theory and Data Compression*, 2¦ ed., Chapman & Hall, ISBN-10: 1584883138, 2003.

[COV06] Cover, T.; Thomas, J.A., *Elements of Information Theory*, 2¦ ed., Wiley-Inter Science, ISBN-10: 0471241954, 2006.

1.4. Problemas

Problema 1

Sean $F_1 = \{1, 2, 3, 4\}$ y $F_2 = \{2, 4, 6, 8\}$ dos fuentes equiprobables independientes. Sea una fuente (F) cuya salida es el mínimo común múltiplo de la salida de las fuentes anteriores $F = mcm(F_1, F_2)$.

a) Calcule la entropía de la fuente $H(F)$.

b) Calcule la información mutua $I(F, F_1)$.

c) Calcule la longitud media de una codificación de Huffman de la fuente F.

d) Suponga que le proponen adivinar F, y como ayuda le dejan escoger entre conocer F_1 o conocer F_2. Qué opción preferiría? Justifique la respuesta y calcule la probabilidad de adivinar F con la opción que ha escogido anteriormente.

Solución

a) A continuación se muestra una tabla con los resultados de aplicar el *mcm*:

F_1	F_2	F
1	2	2
2	2	2
3	2	6
4	2	4
1	4	4
2	4	4
3	4	12
4	4	4
1	6	6
2	6	6
3	6	6
4	6	12
1	8	8
2	8	8
3	8	24
4	8	8

Tabla 1.1: Generación de símbolos de la fuente F

$F = \{2, 4, 6, 8, 12, 24\}$ con las siguientes probabilidades:

$$P(2)=\frac{1}{8},\quad P(4)=\frac{1}{4},\quad P(6)=\frac{1}{4},\quad P(8)=\frac{3}{16},\quad P(12)=\frac{1}{8},\quad P(24)=\frac{1}{16}$$

$H(F)=2.453$ bits

b) $I(F,F_1)=H(F)-H(F|F_1)$

Calculamos $H(F|F_1)$ para todos los valores F_1 y promediamos

$$F_1=\begin{cases}1\Rightarrow H(F|1)=2\\ 2\Rightarrow H(F|2)=2\\ 3\Rightarrow H(F|3)=\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2+\frac{1}{4}\cdot 2=1.5\\ 4\Rightarrow H(F|4)=1.5\end{cases}$$

$H(F|F_1)=\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1.5=1.75$ bits/símbolo

$I(F,F_1)=H(F)-H(F|F_1)=2.453-1.75=0.703$ bits/símbolo

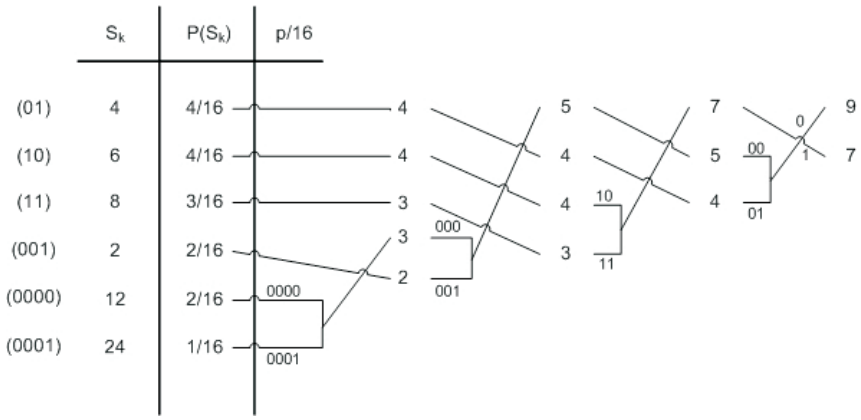


Fig. 1.1: Codificació de Huffman

c) Calculamos una codificación de Huffman de F.

Promediamos la longitud de codificación de cada símbolo por su probabilidad:

$$\bar{l}=\sum_k p(s_k)l_k=\frac{11}{16}\cdot 2+\frac{2}{16}\cdot 3+\frac{3}{16}\cdot 4=\frac{40}{16}=2{,}5 \text{ bits}$$

d) En el apartado *b*, hemos calculado $H(F|F_1)=1.75$. Aquí calculamos $H(F|F_2)$.

$$F_2 = \begin{cases} 2 \Rightarrow H(F|2) = 1.5 \\ 4 \Rightarrow H(F|4) = \frac{3}{4} \cdot \log_2\left(\frac{4}{3}\right) + \frac{1}{4} \cdot 2 = 0.811 \\ 6 \Rightarrow H(F|6) = 0.811 \\ 8 \Rightarrow H(F|8) = 0.811 \end{cases}$$

$H(F|F_2) = 0.98325$ bits/símbolo Puesto que $H(F|F_2) < H(F|F_1)$, es más fácil adivinar F conociendo F_2 (hay menor incertidumbre). En la tabla 1.2 se calcula la probabilidad de adivinar F, si se conoce F_2. Listamos el valor de F que elegimos como más probable para cada uno de los posibles valores de F_2, así como la probabilidad de adivinar correspondiente.

Elijo			p(adivinar)
$F_2 = 2$	$\longrightarrow$	$F = 2$	1/2
$F_2 = 4$	$\longrightarrow$	$F = 4$	3/4
$F_2 = 6$	$\longrightarrow$	$F = 6$	3/4
$F_2 = 8$	$\longrightarrow$	$F = 8$	3/4

Tabla 1.2: Valores de F seleccionados en función de los valores de F_2

Finalmente, ponderando por las distintas probabilidades:

$$p\,(\text{adivinar}) = \frac{1}{4} \cdot \frac{2}{4} + \frac{3}{4} \cdot \frac{3}{4} = \frac{11}{16} = 0.6875$$

Problema 2

Dos fuentes de información, S_1 y S_2, emiten símbolos de un alfabeto $\{A, B, C, D, E, F, G, H, I\}$ con una probabilidad

$$P(A) = 1/3;\ P(B) = P(C) = P(D) = P(E) = P(F) = 1/9; P(G) = P(H) = P(I) = 1/27$$

Ambas fuentes emplean respectivamente un canal de comunicaciones ternario para transmitir la información. Para maximizar la explotación del ancho de banda del canal se emplea en cada caso un codificador de fuente cuyos códigos emplean los símbolos del alfabeto **{–1, 0, 1}**

a) Determine si existe un código instantáneo en el que la codificación de todos los símbolos de fuente dé lugar a palabras código de longitud 2.

b) ¿Cuál es la longitud media mínima de las palabras código para una fuente, S_1 o S_2?

c) Calcule, mediante el algoritmo de Huffman, las palabras código para cada uno de los símbolos de fuente. ¿Cuál es la eficiencia del código resultante?

d) ¿Cuál sería una cota superior de la entropía conjunta de las fuentes S_1 y S_2 en bits?

Se observa en la generación de símbolos de las fuentes que existe una dependencia entre las fuentes S_1 y S_2. Esta dependencia se manifiesta de la siguiente manera:

I) Cuando S_1 emite A, entonces S_2 sólo emite A.
II) Cuando S_1 emite B, C o D, entonces S_2 sólo emite B, C o D.
III) Cuando S_1 emite E, F, G, H o I, entonces S_2 sólo emite E, F, G, H o I.

e) Teniendo en cuenta la dependencia entre las fuentes, calcule la entropía de la fuente S_2 en bits para los casos en los que la fuente S_1 toma el valor: $S_1 = A$ y $S_1 = C$.

Solución

Dos fuentes de información, S_1 y S_2

$$P(A) = \frac{1}{3}; \quad P(B) = P(C) = P(D) = P(E) = P(F) = \frac{1}{9}; \quad P(G) = P(H) = P(I) = \frac{1}{27}$$

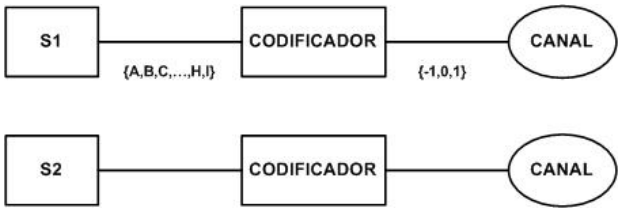


Fig. 1.2: Esquema de codificación

a) Para que un código sea instantáneo, ha de cumplir la desigualdad de Kraft:

$$\sum_{k=1}^{n} D^{-L_k} \leq 1$$

En nuestro caso: $n = 9$ número de símbolos de fuente
$D = 3$ número de símbolos que emplean los códigos
$L_k = 2 \; \forall k$ longitud de todos los códigos

$$\sum_{k=1}^{9} 3^{-2} = \sum_{k=1}^{9} \frac{1}{9} = 1 \leq 1$$

Por tanto, existe un código instantáneo cuyas palabras son de longitud 2.

b) La longitud mínima está determinada por la entropía de la fuente, $\overline{L}_{\min} = H$

Puesto que el código es ternario, debemos utilizar base 3.

$$\overline{L}_{\min} = \sum_{k=1}^{9} p_k \cdot \log_3 \frac{1}{p_k} = -\sum_{k=1}^{9} p_k \cdot \log_3 p_k$$

$$= \frac{1}{3} \cdot \log_3 3 + 5 \cdot \frac{1}{9} \log_3 3^2 + 3 \cdot \frac{1}{27} \log_3 3^3$$

$$= \frac{1}{3} + \frac{10}{9} + \frac{1}{3} = 1{,}77 \text{ dígitos ternarios}$$

$\overline{L}_{\min} = 1{,}77$ dígitos ternarios

c) Calculamos la codificación por Huffman.

Ordenamos por probabilidad	Fuente reducida 1 ordenada	Fuente reducida 2	Fuente reducida 2 ordenada	Fuente reducida 3 ordenada
$A \rightarrow 1/3$	$A \rightarrow 1/3$	$A \rightarrow 1/3$	$A \rightarrow 1/3$	$A \rightarrow 1/3$
$B \rightarrow 1/9$	$B \rightarrow 1/9$	$B \rightarrow 1/9$	$K \rightarrow 1/3$	$K \rightarrow 1/3$
$C \rightarrow 1/9$	$C \rightarrow 1/9$	$C \rightarrow 1/9$	$B \rightarrow 1/9$	$L \rightarrow 1/3$
$D \rightarrow 1/9$	$D \rightarrow 1/9$	$D \rightarrow 1/9$	$C \rightarrow 1/9$	
$E \rightarrow 1/9$	$E \rightarrow 1/9$	$K \rightarrow 1/3$	$D \rightarrow 1/9$	
$F \rightarrow 1/9$	$F \rightarrow 1/9$			
$G \rightarrow 1/27$	$J \rightarrow 1/9$			
$H \rightarrow 1/27$				
$I \rightarrow 1/27$				

Y obtenemos como resultado:

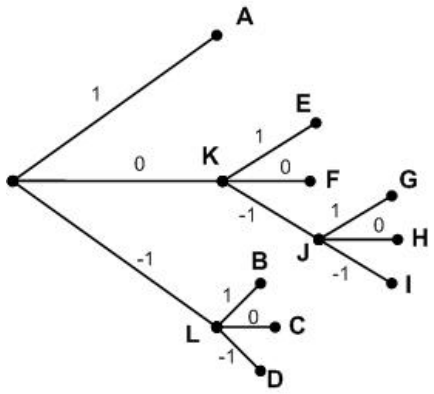


Fig. 1.3: Codificación ternaria de Huffman

A	$\rightarrow$	1		
B	$\rightarrow$	-1	1	
C	$\rightarrow$	-1	0	
D	$\rightarrow$	-1	-1	
E	$\rightarrow$	0	1	
F	$\rightarrow$	0	0	
G	$\rightarrow$	0	-1	1
H	$\rightarrow$	0	-1	0
I	$\rightarrow$	0	-1	-1

Tabla 1.3: Tabla de codificación ternaria

Dado que la longitud del código de cada símbolo coincide con la información que proporciona (en base 3), entonces es inmediato que: $\overline{L}_{\min} = 1{,}77 \implies E = \dfrac{H}{\overline{L}} = 1$

d) Una cota superior de $H(S_1, S_2)$ se obtiene cuando ambas fuentes son independientes:

$H(S_1, S_2) \leq H(S_1) + H(S_2)|_{\{H(S_1)=H(S_2)\}} = 2 \cdot H(S_1) = 2 \cdot H(S_2)$ Para expresar la información en bits, empleamos base 2:

$$H(S_1) = \frac{1}{3} \cdot \log_2 3 + \frac{5}{9} \log_2 3^2 + \frac{3}{27} \log_2 3^3 = 2{,}81 \text{ bits}$$

$$H(S_1, S_2) \leq 5{,}63 \text{ bits}$$

e) Cuando

$S_1 = A \implies S_2 = A$
$S_1 = C \implies S_2 = B$ o $S_2 = C$ o $S_2 = D$

las probabilidades condicionadas son: $P(S_2 = A|S_1 = A) = 1$

Para el caso $S_1 = C$, hay tres símbolos. Considerando que estos símbolos mantienen la relación de probabilidades de la fuente S_2, entonces:

$$P(S_2 = B|S_1 = C) + P(S_2 = C|S_1 = C) + P(S_2 = D|S_1 = C) = 1$$

$$P(S_2 = B|S_1 = C) = \frac{P(S_2 = B)}{P(S_2 = B) + P(S_2 = C) + P(S_2 = D)} = \frac{1}{3}$$

De la misma manera, $P(S_2 = C|S_1 = C) = P(S_2 = D|S_1 = C) = \dfrac{1}{3}$.

Finalmente,

$$H(S_2|S_1 = A) = P(S_2 = A|S_1 = A) \cdot \log_2 \frac{1}{P(S_2 = A|S_1 = A)} = 0$$

$$\begin{aligned} \mathrm{H}(S_2|S_1 = C) &= 3 \cdot P(S_2 = B|S_1 = C) \cdot \log_2 \frac{1}{P(S_2 = B|S_1 = C)} = 3 \cdot \frac{1}{3} \log_2 3 \\ &= 1{,}58 \text{ bits} \end{aligned}$$

Se puede comprobar que $H(S_2|S_1) = 1{,}23$ bits

$H(S_1, S_2) = H(S_1) + H(S_2|S_1) = 5{,}1$ bits

Problema 3

Una fuente binaria simétrica F emite ráfagas de longitud L, con $L > 0$, según una distribución geométrica de parámetro p:

$$Prob[L = k] = p^{k-1}(1-p), \quad con \quad k = 1, 2, \ldots \text{y} \quad 0 < p < 1$$

a) Proponga un modelo markoviano de la fuente F, con memoria 1, y evalúe su entropía H(F) para un valor p genérico. Particularice el resultado para $p = 1/2$.

b) Aplicando una codificación de fuente por ráfagas, resulta una fuente F' cuyos símbolos representan la longitud de las ráfagas de F, $\{1, 2, 3, \ldots\}$.

I) Determine la entropía $H(F')$ para $p = 1/2$.

II) Suponiendo que, en la práctica, la fuente no genera ráfagas de longitud mayor a 7, y despreciando la probabilidad de estos casos, realice una codificación binaria de Huffman de F' para el caso $p = 1/2$.

III) A partir de los resultados obtenidos en los apartados anteriores, analice las ventajas y los inconvenientes de la codificación por ráfagas para el caso $p = 1/2$.

Solución

a) La fuente F genera ráfagas de 0 y 1. Un modelo markoviano con memoria 1 de F sería:

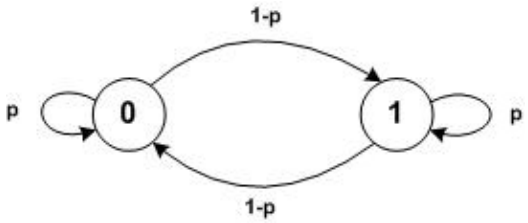


Fig. 1.4: Fuente markoviana de memoria 1

Por simetría: $Prob(F=0) = Prob(F=1) = \frac{1}{2}$

Por simetría: $H(F_n|F_{n-1}=1) = H(F_n|F_{n-1}=0) = H(p)$

Luego: $H(F) = P(F=1)\cdot H(F_n|F_{n-1}=1) + P(F=0)\cdot H(F_n|F_{n-1}=0) = H(p)$

Si $p = 1/2 \Rightarrow H(F) = 1$ bit/símbolo y la fuente no tiene memoria.

b) $F' = \{1, 2, 3, 4, \ldots\}$ $Prob(F'=k) = Prob[L=k] = p^{k-1}(1-p)$

I) $H(F') = p_1 \log_2 \frac{1}{p_1} + p_2 \log_2 \frac{1}{p_2} + p_3 \log_2 \frac{1}{p_3} + \cdots$

$$\begin{aligned} p_1 &= Prob[F'=1] = 1-p \\ p_2 &= Prob[F'=2] = (1-p)\,p \\ p_3 &= Prob[F'=3] = (1-p)^2 \\ \vdots &= \vdots \\ p_k &= Prob[F'=K] = (1-p)\,p^{k-1} \end{aligned}$$

$$H(F') = \sum_{k=1}^{\infty} p_k \log_2 \frac{1}{p_k} = \sum_{k=1}^{\infty} \left[(1-p)\,p^{k-1} \log_2 \frac{1}{(1-p)\,p^{k-1}} \right]$$

$$H(F') = (1-p) \sum_{k=1}^{\infty} -p^{k-1} \left[\log_2 (1-p) + \log_2 p^{k-1} \right]$$

$$H(F') = -(1-p) \left[\sum_{k=1}^{\infty} \log_2 (1-p)\,p^{k-1} + \sum_{k=1}^{\infty} (k-1)\,p^{k-1} \log_2 p \right]$$

$$= -(1-p) \log_2 (1-p) \sum_{k=1}^{\infty} p^{k-1} - (1-p) \log_2 p \sum_{k=1}^{\infty} (k-1)\,p^{k-1}$$

Dado que

$$\sum_{0}^{\infty} p^k = \frac{1}{1-p} \quad \text{y} \quad \sum_{0}^{\infty} k\,p^k = \frac{p}{(1-p)^2}$$

se obtiene:

$$H(F') = -\frac{p}{1-p} \log_2 p - \log_2 (1-p)$$

Si $p = 1/2$, entonces

$$H(F') = -\log_2 \frac{1}{2} - \log_2 \frac{1}{2} = 2 \text{ bits } F'$$

II) $K = \{1, 2, 3, 4, 5, 6, 7\}$ $Prob[L \geq 8] \simeq 0$

Suponiendo $p = 1/2$, la codificación de Huffman será

$p_1 = 0.5$
$p_2 = 0.250$
$p_3 = 0.125$
$p_4 = 0.0625$
$p_5 = 0.03125$
$p_6 = 0.015625$
$p_7 = 0.0078125$

Puesto que las probabilidades difieren tanto entre sí, la codificación es directa.

En particular, se tiene que:

$p_A = (p_6 + p_7) = 0.0234375$
$p_B = (p_5 + p_A) = 0.0546$
$p_C = (p_4 + p_B) = 0.11718$
$p_D = (p_3 + p_C) = 0.242$
$p_E = (p_2 + p_D) = 0.492$

Desarrollando la codificación de Huffman obtenida, tenemos:

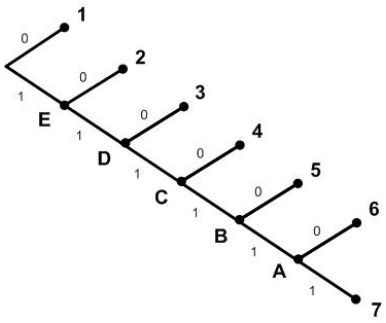


Fig. 1.5: Codificación de Huffman de la longitud de las ráfagas

y la tabla de codificación resultante es:

$$
\begin{aligned}
1 &\rightarrow 0 \\
2 &\rightarrow 1\,0 \\
3 &\rightarrow 1\,1\,0 \\
4 &\rightarrow 1\,1\,1\,0 \\
5 &\rightarrow 1\,1\,1\,1\,0 \\
6 &\rightarrow 1\,1\,1\,1\,1\,0 \\
7 &\rightarrow 1\,1\,1\,1\,1\,1
\end{aligned}
$$

III) Obsérvese que, de no haberse truncado la longitud de las ráfagas, cada valor de longitud requiere un nṁero de dígitos de codificación igual al tamaño de la ráfaga. Por tanto, esta codificación no ofrece ninguna ventaja respecto a enviar el valor de cada símbolo de F para $p = 1/2$. De forma equivalente, si $p = 1/2$, la fuente no tiene memoria y los símbolos son equiprobables, con lo que la codificación a ráfagas no tiene ventajas.

Problema 4

La trayectoria de un coche se puede modelar como la de un objeto que se mueve a través de una retícula cuadriculada con pasos elementales, en direcciones verticales u horizontales, dando un único paso cada vez. Así, se puede representar su movimiento como una sucesión de símbolos del conjunto N, S, E, y W, que representan los sucesivos pasos en las direcciones norte, sur, este y oeste, respectivamente.

El comportamiento de este coche tiene memoria: el 50 % de las ocasiones repite el movimiento anterior y, en el resto de los casos, da un giro de 90° a la derecha (con probabilidad 30 %) o a la izquierda (con probabilidad del 20 %) respecto del paso anterior.

Se pide:

a) Modelar el proceso que describe el movimiento.
b) Calcular la probabilidad de cada uno de los símbolos.
c) Calcular la tasa de entropía de esta fuente de información.
d) Diseñar un codificador Huffman de esta fuente.

Antes \ Ahora	N	S	E	W
N	0,5	–	0,3	0,2
S	–	0,5	0,2	0,3
E	0,2	0,3	0,5	–
W	0,3	0,2	–	0,5

Solución

a) Se puede modelar el proceso con una cadena de Markov, memoria 1, cuyo diagrama de probabilidades de transmisión de estados es:

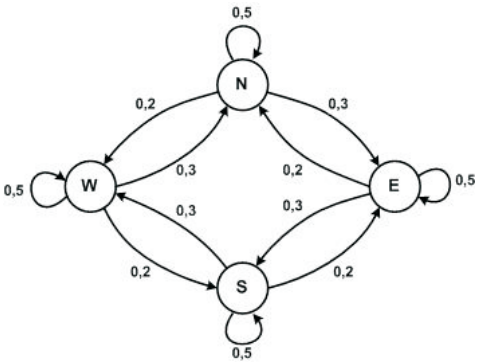


Fig. 1.6: Diagrama de estados de proceso de movimiento del vehículo

b)

$$
\begin{aligned}
P(N) &= P(N|N)\cdot P(N) + P(N|S)\cdot P(S) + P(N|E)\cdot P(E) + P(N|W)\cdot P(W)\\
P(S) &= P(S|N)\cdot P(N) + P(S|S)\cdot P(S) + P(S|E)\cdot P(E) + P(S|W)\cdot P(W)\\
P(E) &= P(E|N)\cdot P(N) + P(E|S)\cdot P(S) + P(E|E)\cdot P(E) + P(E|W)\cdot P(W)\\
P(W) &= P(W|N)\cdot P(N) + P(W|S)\cdot P(S) + P(W|E)\cdot P(E) + P(W|W)\cdot P(W)
\end{aligned}
$$

$$
\begin{aligned}
0.5\cdot P(N) &= 0.2\cdot P(E) + 0.3\cdot P(W) \longrightarrow P(N) = 0.4\cdot P(E) + 0.6\cdot P(W)\\
0.5\cdot P(S) &= 0.3\cdot P(E) + 0.2\cdot P(W) \longrightarrow 2P(S) = 0.6\cdot P(E) + 0.4\cdot P(W)\\
0.5\cdot P(E) &= 0.3\cdot P(N) + 0.2\cdot P(S) \longrightarrow 2P(E) = 0.6\cdot P(N) + 0.4\cdot P(S)\\
0.5\cdot P(W) &= 0.2\cdot P(N) + 0.3\cdot P(S) \longrightarrow 2P(W) = 0.4\cdot P(N) + 0.6\cdot P(S)
\end{aligned}
$$

Resolviendo:

$$P(N) + P(S) = P(E) + P(W)$$

$$P(N) + P(S) + P(E) + P(W) = 1 \longrightarrow 2\cdot P(N) + 2\cdot P(S) = 1$$

$$P(N) + P(S) = \frac{1}{2}$$

$$P(E) + P(W) = \frac{1}{2}$$

$$
\begin{aligned}
P(N) &= 0.4\cdot (\overbrace{0.6\cdot P(N) + 0.4\cdot P(S)}^{P(E)}) + 0.6\cdot (\overbrace{0.4\cdot P(N) + 0.6\cdot P(S)}^{P(W)})\\
&= 0.24\cdot P(N) + 0.16\cdot P(S) + 0.24\cdot P(N) + 0.36\cdot P(S)\\
&= 0.48\cdot P(N) + 0.52\cdot P(S)
\end{aligned}
$$

$$0.52\cdot P(N) = 0.52\cdot P(S) \implies P(N) = P(S) \implies P(N) = P(S) = \frac{1}{4}$$

$$P(E) = 0.6 \cdot \overbrace{(0.4 \cdot P(E) + 0.6 \cdot P(W))}^{P(N)} + 0.4 \cdot \overbrace{(0.6 \cdot P(E) + 0.4 \cdot P(W))}^{P(S)}$$

$$= 0.24 \cdot P(E) + 0.36 \cdot P(W) + 0.24 \cdot P(E) + 0.16 \cdot P(W)$$

$$P(E) = P(W) \implies P(E) = P(W) = \frac{1}{4}$$

c) $H(F) = H(F|N) \cdot P(N) + H(F|S) \cdot P(S) + H(F|E) \cdot P(E) + H(F|W) \cdot P(W)$

Todas las entropías condicionadas son iguales, dado que las probabilidades de transición en todos los estados son iguales.

$$\begin{aligned} H(F|N) &= H(F|S) = H(F|E) = H(F|W) \\ H(F|N) &= P(N|N) \cdot \log_2\left(\frac{1}{P(N|N)}\right) + P(S|N) \cdot \log_2\left(\frac{1}{P(S|N)}\right) + \\ &+ P(E|N) \cdot \log_2\left(\frac{1}{P(E|N)}\right) + P(W|N) \cdot \log_2\left(\frac{1}{P(W|N)}\right) \\ &= 0.5 \cdot \log_2\left(\frac{1}{0,5}\right) + 0.3 \cdot \log_2\left(\frac{1}{0.3}\right) + 0.2 \cdot \log_2\left(\frac{1}{0.2}\right) \\ &= 0.5 + 0.5211 + 0.4644 = 1.4855 \text{ [bits} \end{aligned}$$

d) El código Huffman no prevé que la fuente tenga memoria:

N 1/4	A 1/2	A 1/2	N 01
S 1/4	N 1/4	B 1/2	S 00
E 1/4	S 1/4		E 11
W 1/4			W 10

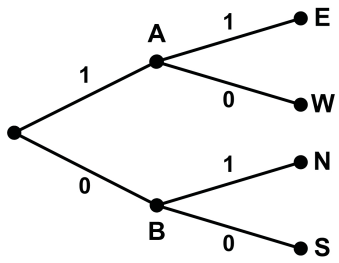


Fig. 1.7: Codificación binaria de Huffman

Problema 5

Una fuente binaria con memoria 1 envía, de forma periódica, símbolos a un codificador de fuente cada T_F.

El codificador aplica una extensión de fuente concatenando dichos símbolos de dos en dos, de forma que trabaja con un alfabeto $\{X, \overline{X}, Y, \overline{Y}\}$. El comportamiento de la fuente extendida puede ser modelado mediante la cadena de Markov:

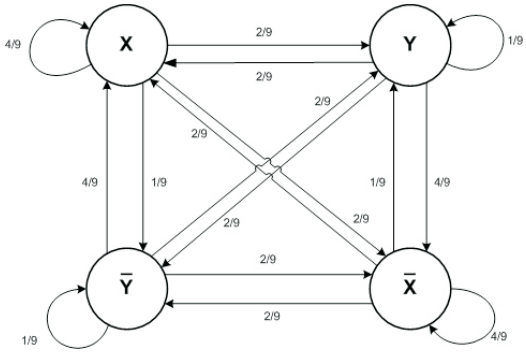


Fig. 1.8: Cadena de Markov de la fuente extendida

a) Para el régimen estacionario, calcule la probabilidad de que la fuente extendida genere cada uno de los símbolos. Tenga en cuenta las simetrías de la cadena de Markov para el cálculo.

b) Determine la entropía de la fuente extendida en bits, H(Fe).

c) Suponiendo que la codificación de la fuente extendida obtiene una longitud media de 1,88 dígitos binarios por símbolo, halle el valor mínimo de T_F para un canal de 64 Kbps.

d) Teniendo en cuenta que la fuente binaria se puede modelar con la cadena de Markov:

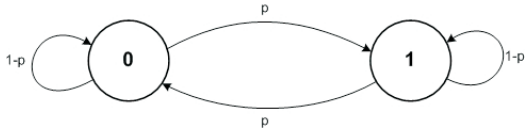


Fig. 1.9: Cadena de Markov de la fuente binaria

Identifique el valor de p a partir del modelo de fuente extendida y la asociación entre los valores del alfabeto de la fuente extendida y los pares de símbolos binarios (tenga en cuenta que los valores $X - \bar{X}$ e $Y - \bar{Y}$ son complementarios entre sí).

e) Para un valor de $p = 1/3$, halle la relación entre entropías de la fuente extendida y la fuente binaria. Discuta los valores obtenidos con respecto al caso sin memoria.

Solución

a) Se observa que $P(X) = P(\overline{X})$, $P(Y) = P(\overline{Y})$

Se debe cumplir $P(X) + P(\overline{X}) + P(Y) + P(\overline{Y}) = 1$ de lo cual se deriva que $P(X) + P(Y) = 1/2$

Por tanto, falta una ecuación, que se deriva de la cadena de Markov. Por ejemplo, para el estado X, se verifica en régimen estacionario que

$$P(X) \cdot \left[\frac{2}{9} + \frac{2}{9} + \frac{1}{9}\right] = P(Y) \cdot \frac{2}{9} + P(\overline{Y}) \cdot \frac{4}{9} + P(\overline{X}) \cdot \frac{2}{9}$$

Simplificando:

$$5P(X) = 2P(Y) + 4P(\overline{Y}) + 2P(\overline{X}) \quad \Rightarrow \quad P(X) = 2P(Y)$$

Resolviendo:

$$P(X) = \frac{1}{3}, \quad P(\overline{X}) = \frac{1}{3}, \quad P(Y) = \frac{1}{6}, \quad P(\overline{Y}) = \frac{1}{6}$$

b) $H(F_e) = P(X) \cdot H(F_e|X) + P(\overline{X}) \cdot H(F_e|\overline{X}) + P(Y) \cdot H(F_e|Y) + P(\overline{Y})H(F_e|\overline{Y})$

$$H(F_e|X) = P_{X/X} \log_2 \frac{1}{P_{X/X}} + P_{\overline{X}/X} \log_2 \frac{1}{P_{\overline{X}/X}} + P_{Y/X} \log_2 \frac{1}{P_{Y/X}} + P_{\overline{Y}/X} \log_2 \frac{1}{P_{\overline{Y}/X}}$$

$$H(F_e|X) = \frac{4}{9}\log_2 \frac{9}{4} + \frac{2}{9}\log_2 \frac{9}{2} + \frac{2}{9}\log_2 \frac{9}{2} + \frac{1}{9}\log_2 9$$

$$H(F_e|X) = \frac{4}{9}\log_2 \frac{9}{4} + \frac{4}{9}\log_2 \frac{9}{2} + \frac{1}{9}\log_2 9$$

$$H(F_e|X) = \frac{4}{9}(\log_2 9 - 2) + \frac{4}{9}(\log_2 9 - 1) + \frac{1}{9}\log_2 9$$

$$H(F_e|X) = \log_2 9 - \frac{12}{9} = 2\log_2 3 - \frac{4}{3} = 1{,}83 \text{ bits}$$

Dado que todos los estados tienen el mismo conjunto de probabilidades de transición,

$$H(F_e|X) = H(F_e|\overline{X}) = H(F_e|Y) = H(F_e|\overline{Y})$$

por lo que:

$$H(F) = H(F_e|X) = 2\log_2 3 - \frac{4}{3} = 1{,}83 \text{ bits}$$

c) $L_{F_e} = 1{,}88$ dig bin/sim F_e

La velocidad máxima de la fuente extendida será:

$$v_{F_e} = \frac{C}{L_{F_e}} = \frac{64.000}{1.88} = 34042{,}55 \text{ sim } F_e/\text{s}$$

La velocidad máxima de la fuente será:

$v_F = 2 \cdot v_{F_e} = 68.085{,}10$ sim F/s

$T_F = \frac{1}{v_F} = 1.468 \cdot 10^{-5} = 14.68\ \mu$s (tiempo medio mínimo entre símbolos)

d) El alfabeto extendido será $\{00, 01, 10, 11\}$.

La fuente extendida se encuentra en el estado 00 cuando las dos últimas generaciones de la fuente elemental han sido 0. Esto implica que, cuando se realiza una transición desde el estado 00 de la fuente extendida, la fuente elemental se encuentra en el estado 0 y realiza dos nuevas generaciones. Para este caso, las cuatro posibilidades son:

I) $X = 00 \longrightarrow X = 00$: ocurre cuando la fuente elemental desde el estado 00 realiza dos generaciones de 0, cuya probabilidad es $(1-p)^2$.

II) $X = 00 \longrightarrow X = 01$: ocurre cuando la fuente elemental desde el estado 0 genera otro 0 y luego un 1, cuya probabilidad es $(1-p)\,p$.

III) $X = 00 \longrightarrow X = 10$: ocurre cuando la fuente elemental desde el estado 0 genera un 1 y luego vuelve a generar otro 0, cuya probabilidad es p^2.

IV) $X = 00 \longrightarrow X = 11$: ocurre cuando la fuente elemental desde el estado 0 genera dos 1 repetidamente, con probabilidad $p\,(1-p)$.

Gráficamente, se representan las transiciones desde el estado 00.

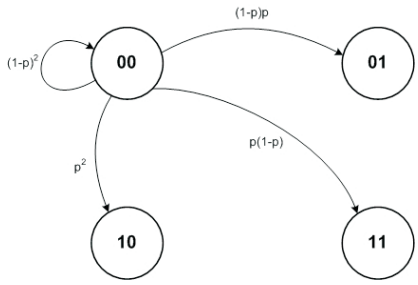


Fig. 1.10: Diagrama de transición de estados

Dadas las simetrías del modelo de la fuente extendida, podemos identificar dos soluciones distintas:

$$\text{I)}\ X = 00,\ \overline{X} = 11,\ Y = 10,\ \overline{Y} = 01 \quad \Rightarrow \quad \begin{cases} p(1-p) = \dfrac{2}{9} \\ p^2 = \dfrac{1}{9} \end{cases} \quad \Rightarrow \quad p = 1/3$$

$$\text{II)}\ X = 01,\ \overline{X} = 10,\ Y = 00,\ \overline{Y} = 11 \Rightarrow \begin{cases} p(1-p) = \dfrac{2}{9} \\ p^2 = \dfrac{4}{9} \end{cases} \quad \Rightarrow \quad p = 2/3$$

e) $p = 1/3$

$$H(F) = P(0) \cdot H(F|0) + P(1) \cdot H(F|1)$$

$$P(0) = P(1) = 1/2 \text{ por simetría}$$

$$H(F|0) = H(p) = \frac{1}{3}\log_2 3 + \frac{2}{3}\log_2 \frac{3}{2}$$

$$H(F|0) = \frac{1}{3}\log_2 3 + \frac{2}{3}\log_2 3 - \frac{2}{3}\log_2 2$$

$$H(F|0) = \log_2 3 - \frac{2}{3}$$

La probabilidad de transición es la misma para el estado 1, por lo que

$$SH(F) = H(F|0) = \log_2 3 - \frac{2}{3} = 0{,}9183 \text{ bits}$$

La solución será:

$$\frac{H(F_e)}{H(F)} = \frac{2\log_2 3 - \dfrac{4}{3}}{\log_2 3 - \dfrac{2}{3}} = 2 \text{ (concatenación de 2 símbolos)}$$

Si F no tuviera memoria, $p = 1/2$, la fuente extendida sería la agrupación de símbolos independientes de F, por lo que la entropía crece linealmente con el número de símbolos concatenados. Así:

$$H(F_e) = 2H(F)$$

Por tanto, se mantiene la misma relación, como cabía esperar, puesto que la fuente extendida no añade ningún desorden.

Problema 6

Se desea realizar la compresión del un fichero cuyo contenido es:

A B D B D A D C A C C A D C B B

Suponiendo que se ha fijado a priori para cada símbolo de la fuente la siguiente asociación binaria de dos bits:

$$\{A =' 00',\ B =' 01',\ C =' 10',\ D =' 11'\}$$

a) Indique cuál es la mínima longitud en bits del resultado de la compresión del fichero.

b) Exprese en hexadecimal el resultado de la compresión del fichero cuando:

I) Se emplea el algoritmo LZ-77 con una memoria de almacenamiento de 8 posiciones (3 bits de direccionamiento).
II) Se emplea el algoritmo LZ-78 con un diccionario de 64 posiciones (longitud final igual a 6 bits de direccionamiento).
III) Se emplea el algoritmo LZW con un diccionario de 256 posiciones (8 bits de direccionamiento).

Solución

a) Hay 16 símbolos de fuente en el fichero. Si suponemos que son equiprobables, necesitaríamos $16 \cdot 2$ bits $= 32$ bits. En el fichero aparecen los símbolos con probabilidad 1/4, por lo que son equiprobables. Luego, de igual manera, con la estadística del fichero, necesitaríamos al menos 32 bits.

b) I) LZ-77 $\rightarrow$ (Pos, Long, Carác)

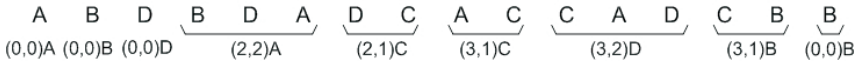


Fig. 1.11: Diagrama de transición de estados

		Pos	Long	Car		
(0, 0) A	$\longrightarrow$	000	000	00	$\longrightarrow$	00h
(0, 0) B	$\longrightarrow$	000	000	01	$\longrightarrow$	01h
(0, 0) D	$\longrightarrow$	000	000	11	$\longrightarrow$	03h
(2, 2) A	$\longrightarrow$	010	010	00	$\longrightarrow$	48h
(2, 1) C	$\longrightarrow$	010	001	10	$\longrightarrow$	46h
(3, 1) C	$\longrightarrow$	011	001	10	$\longrightarrow$	66h
(3, 2) D	$\longrightarrow$	011	010	11	$\longrightarrow$	6Bh
(3, 1) B	$\longrightarrow$	011	001	01	$\longrightarrow$	65h
(0, 0) B	$\longrightarrow$	000	000	01	$\longrightarrow$	01h

II) LZ78 → (Pos, Carác)

A	B	D	B D	A D	C	A C	C A	D C	B B
(0,A)	(0,B)	(0,D)	(2,D)	(1,D)	(0,C)	(1,C)	(6,A)	(3,C)	(2,B)

posición	carácter
000001	A
000010	B
000011	D
000100	BD
000101	AD
000110	C
000111	AC
001000	CA
001001	DC
001010	BB

(0, A)	⟶	00000000	⟶	00h
(0, B)	⟶	00000001	⟶	01h
(0, D)	⟶	00000011	⟶	03h
(2, D)	⟶	00001011	⟶	0Bh
(1, D)	⟶	00000111	⟶	07h
(0, C)	⟶	00000010	⟶	02h
(1, C)	⟶	00000110	⟶	06h
(6, A)	⟶	00011000	⟶	18h
(3, C)	⟶	00001110	⟶	DEh
(2, B)	⟶	00001001	⟶	09h

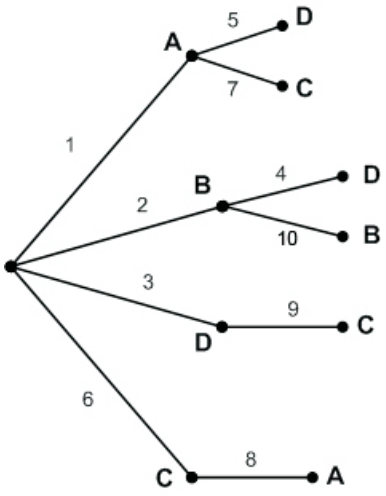


Fig. 1.12: Codificación LZ78

III) LZW

A	B	D	B D	A	D	C	A	C	C A	D C	B	B
0	1	3	5	0	3	2	0	2	10	9	1	1

Diccionario			Codificación		
0	⟶	A	0	⟶	00h
1	⟶	B	1	⟶	01h
2	⟶	C	3	⟶	03h
3	⟶	D	5	⟶	05h
4	⟶	AB	0	⟶	00h
5	⟶	BD	3	⟶	03h
6	⟶	DB	2	⟶	02h
7	⟶	BDA	0	⟶	00h
8	⟶	AD	2	⟶	02h
9	⟶	DC	10	⟶	0Ah
10	⟶	CA	9	⟶	09h
11	⟶	AC	1	⟶	01h
12	⟶	CC	1	⟶	01h
13	⟶	CAD			
14	⟶	DCB			
15	⟶	BB			

Problema 7

Sea una fuente de información sin memoria cuyo alfabeto es de 3 símbolos $\{A, B, C\}$, con $P(A) = 0.5 \qquad P(B) = P(C) = 0.25$

a) Calcule el tiempo mínimo necesario para transmitir 10.000 símbolos de fuente a través de un canal ($W = 3$ KHz) cuya relación señal a ruido a la entrada del receptor es $S/N = 7$ (en escala lineal).

b) Codifique la secuencia *ABACAA* mediante un codificador de Huffman.

c) Codifique la secuencia *ABACAA* mediante un codificador aritmético.

d) Descodifique la secuencia 0011426 mediante un codificador de LZW, con un diccionario cargado inicialmente con *A* en la posición 0, *B* en la 1 y *C* en la 2.

e) ¿Cuál de las codificaciones anteriores considera más apropiada para la fuente en cuestión? Razone la respuesta.

Solución

a) En el mejor de los casos (tiempo de transmisión mínimo): *i*) No podremos transmitir por encima de la capacidad de canal (C). *ii*) Cada símbolo de fuente lo podremos comprimir, por término medio, hasta el valor de la entropía.

$$V_{t_{máx}} = C = W \log_2 \left(1 + \frac{S}{N}\right) = 9.000 \text{ bps} \qquad L_{mín} = 10.000 \cdot H = 15.000 \text{ bits}$$

$$H = \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{4}\log_2 4 = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 1{,}5 \text{ bits}$$

$$t_{mín} = \frac{L_{mín}}{V_{t_{máx}}} = \frac{15.000}{9.000} = \frac{5}{3}\text{s}$$

b) Codificación de Huffman de la secuencia *ABACAA*:

A	0,25	0,5	0
B	0,25	0,5	1
C	0,25		

Por lo que $A \rightarrow 0, B \rightarrow 1\,0, C \rightarrow 1\,1$

Y la secuencia $ABACAA \rightarrow 0\,1\,0\,0\,1\,1\,0\,0$

c) Codificación aritmética de la secuencia:

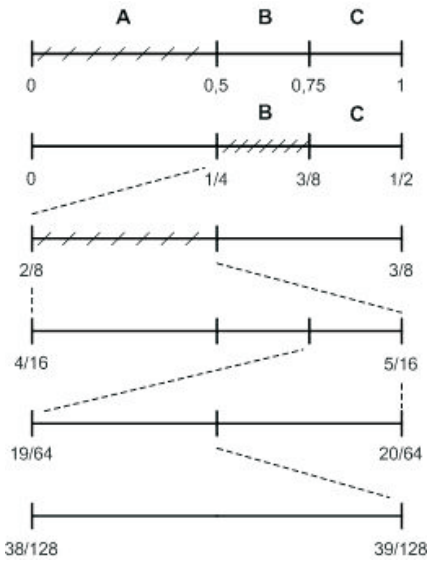


Fig. 1.13: Codificación aritmética

La secuencia $ABACAA \in \left[\frac{38}{128}, \frac{39}{128}\right)$. Por tanto, cualquier punto dentro de este segmento permite codificar esa secuencia de entrada con un único valor.

d)

RX:	0	0	1	1	4	2	6
SALIDA:	*A*	*A*	*B*	*B*	*AB*	*C*	*BA*
AÑADIR DICC:	—	*AA*	*AB*	*BB*	*BA*	*ABC*	*CB*

DIC:							
	0	*A*	3	*AA*	6	*BA*	
	1	*B*	4	*AB*	7	*ABC*	
	2	*C*	5	*BB*	8	*CB*	

SALIDA: *AABBABCBA*

e) En este caso, en la codificación de Huffman, $\bar{l} = H$, por lo que es el más apropiado.

Problema 8

Un sistema de transmisión de datos emplea un codificador de fuente y un cifrador en flujo basado en un simple LFSR. La fuente F que emplea el sistema carece de memoria y emite símbolos del alfabeto $\{A, B\}$ cuyas probabilidades de generación son $p_A = 0.9$ y $p_B = 0.1$. La transmisión se realiza sobre un canal cuya capacidad es de C bps. La codificación binaria aplicada utiliza una extensión de fuente de orden 2 (concatenación de símbolos de 2 en 2) y el algoritmo de Huffman. El cifrador en flujo emite una secuencia cifrante K cuyos valores 1 y 0 son equiprobables. El flujo binario de salida del codificador de fuente se ha denominado X y el entregado al canal Y, resultado de $X \oplus K$.

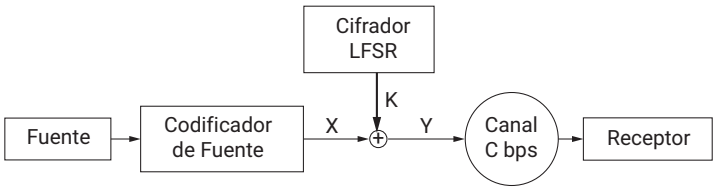


Fig. 1.14: Esquema de transmisión de datos con un cifrador-aleatorizador

a) Determine la entropía de la fuente $H(F)$.
b) Determine la entropía de la fuente extendida $H(F^2)$.
c) Halle la codificación de Huffman de la fuente extendida y calcule la eficiencia resultante E_{F^2}.
d) Para un canal con C = 64 Kbps, determine la máxima velocidad de emisión de símbolos de la fuente extendida por segundo (v_F) que acepta el sistema.
e) Calcule $H(Y|X)$, $H(Y|K)$ y $H(X, Y)$.
f) Determine el valor de la información mutua $I(X; K)$.

Solución

a) Fuente $\{A, B\}$: $p_A = 0.9$; $p_B = 0.1$

$$H(F) = p_A \log \frac{1}{p_A} + p_B \log \frac{1}{p_{AB}} = 0{,}469 \text{ bits } F$$

b) Fuente extendida $F^2 = \{AA, AB, BA, BB\}$

$$H(F^2) = 2 \cdot H(F) = 0{,}938 \text{ bits/ } F^2$$

c) Huffman de F^2

Ord. probabilidad	Fuente reducida 1	Fuente reducida 2
$AA \rightarrow 0.81$	$AA \rightarrow 0.81$	$AA \rightarrow 0.81$
$AB \rightarrow 0.09$	$C \rightarrow 0.1$	$D \rightarrow 0.19$
$BA \rightarrow 0.09$	$AB \rightarrow 0.09$	
$BB \rightarrow 0.01$		

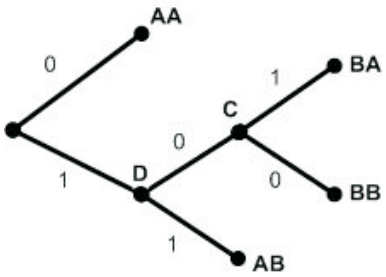


Fig. 1.15: Codificación binaria de Huffman

Codificación: $AA \rightarrow 0$
$AB \rightarrow 1\ 1$
$BA \rightarrow 1\ 0\ 1$
$BB \rightarrow 1\ 0\ 0$

$$L_{F^2} = 0.81 \cdot 1 + 0.09 \cdot 2 + 0.09 \cdot 3 + 0.01 \cdot 3$$
$$= 1.29 \text{ dígitos binarios/símbolo extendido } F^2$$

$$E_{F^2} = \frac{H(F^2)}{L_{F^2}} = \frac{0.938}{1.29} = 0.727$$

Obsérvese que la eficiencia ha aumentado notablemente con respecto a la codificación de la fuente F.

d) T_F = tiempo de símbolo de fuente, $v_F = \dfrac{1}{T_F}$

$$\frac{L_{F^2}}{2 \cdot T_F} \leq C \quad \Longrightarrow \quad \frac{L_{F^2}}{2} \cdot v_F \leq C \quad \Longrightarrow \quad v_{F_{\text{máx}}} = \frac{2 \cdot C}{L_{F^2}} = 99.224{,}8 \text{ símbolos de F/s}$$

Obsérvese que, si se aplica directamente la codificación de Huffman, la velocidad de emisión de la fuente es igual a la capacidad del canal, por tanto 64.000 símbolos/s

e) Entropías, dado que X y K son independientes:

I) $H(Y|X) = H(K) = 1$ bit/símbolo binario X.

II) $H(Y|K) = H(X)$

Para hallar $H(X)$, se puede tener en cuenta que el codificador de fuente no introduce ningún desorden y realiza una transformación reversible. Por tanto, se debe relacionar el número de símbolos de X (dígitos binarios) por símbolo de F que realiza la codificación. Esta relación queda expresada por la longitud media de la codificación:

$L_{F^2} = 1{,}29$ dig bin/símbolo $\quad F^2 = 1{,}29$ simb X/simb F^2

Considerando que el codificador realiza una concatenación de 2 símbolos de F:

$$\begin{aligned} H(X) &= H(F) \cdot \frac{\text{bits}}{\text{sim } F} \cdot \frac{2 \text{ sim } F}{1 \text{ sim } F^2} \cdot \frac{1 \text{ sim } F^2}{L_{F^2} \text{ sim } X} \\ &= 0.469 \cdot 2 \cdot \frac{1}{1.29} \cdot \frac{\text{bits}}{\text{sim } X} = 0{,}722 \text{ bits} \end{aligned}$$

Se observa que $H(X)$ no es más que la eficiencia de la codificación realizada, es decir, el promedio de información transportada por dígito binario. Así:

$$H(X) = E_{F^2} = \lim_{n \to \infty} \frac{nH(F)}{\frac{1}{2}L_{F^2}} = \frac{2H(F)}{L_{F^2}}$$

III) $H(X;Y) = H(X) + H(Y|X) = H(X) + H(K)$

En las mismas unidades, se suma:

$H(X,Y) = H(X) + H(K) = 0.72 + 1 = 1{,}72$ bits

f) Información mutua

$I(X,K) = H(X) - H(X|K) = 0$ (independientes)

Problema 9

Un sistema de transmisión de datos está compuesto por una fuente binaria X y un canal binario con borrados, cuya salida denominaremos Y. La fuente emite el símbolo 0 con probabilidad α y el símbolo 1 con probabilidad 1-α. El canal se caracteriza por la matriz estocástica:

$$Q = \begin{pmatrix} 1-\rho & \rho & 0 \\ 0 & \rho & 1-\rho \end{pmatrix}$$

donde ρ es la probabilidad de recibir un borrado (B) a la salida del canal cuando se emite un símbolo binario (0, 1).

a) Halle la relación entre $H(Y)$ y $H(X)$. Razone el resultado obtenido para los casos $\rho = 0$ y $\rho = 1$.
b) Calcule $H(X|Y)$.
c) Determine $H(Y|X)$.
d) Indique cuál es el valor de la información mutua $I(X;Y)$.
e) Especifique cuál es el valor de la capacidad C del canal con borrados en bits por símbolo.

Nota: Intente expresar los resultados utilizando $H(\alpha)$ y $H(\rho)$.

Solución

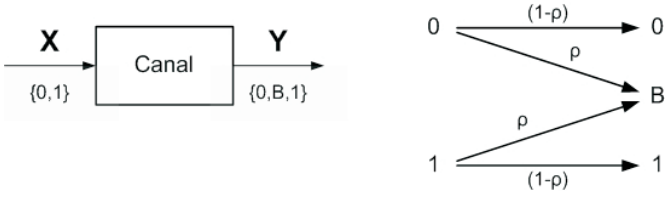


Fig. 1.16: Esquema de transmisión de datos sobre el canal con borrado

a) $H(X) = \alpha \cdot \log_2\left(\frac{1}{\alpha}\right) + (1-\alpha) \cdot \log_2\left(\frac{1}{1-\alpha}\right) = H(\alpha)$

$H(Y) = \sum_i P(Y=i) \cdot \log_2\left(\frac{1}{P(Y=i)}\right)$, donde $i \in \{0, B, 1\}$

$P(Y=0) = (1-\rho) \cdot P(X=0) + 0 \cdot P(X=1) = (1-\rho) \cdot \alpha$

$P(Y=B) = \rho \cdot P(X=0) + \rho \cdot P(X=1) = \rho$

$P(Y=1) = 0 \cdot P(X=0) + (1-\rho) \cdot P(X=1) = (1-\rho) \cdot (1-\alpha)$

$$H(Y) = (1-\rho)\cdot\alpha\cdot\log_2\frac{1}{(1-\rho)\cdot\alpha}+\rho\cdot\log_2\frac{1}{\rho}+(1-\rho)\cdot(1-\alpha)\cdot\log_2\frac{1}{(1-\rho)\cdot(1-\alpha)}$$

$$\begin{aligned}H(Y) =&(1-\rho)\cdot\alpha\cdot\left[\log_2\frac{1}{1-\rho}+\log_2\frac{1}{\alpha}\right]+\rho\cdot\log_2\frac{1}{\rho}+(1-\rho)\cdot(1-\alpha)\cdot\\&\cdot\left[\log_2\frac{1}{1-\rho}+\log_2\frac{1}{1-\alpha}\right]=\\&=\alpha\cdot\left[(1-\rho)\log_2\frac{1}{1-\rho}\right]+(1-\rho)\cdot\left[\alpha\cdot\log_2\frac{1}{\alpha}\right]+\rho\cdot\log_2\frac{1}{\rho}+\\&+(1-\alpha)\cdot\left[(1-\rho)\cdot\log_2\frac{1}{1-\rho}\right]+(1-\rho)\cdot\left[(1-\alpha)\cdot\log_2\frac{1}{1-\alpha}\right]\\&=H(\rho)+(1-\rho)\cdot H(\alpha)\end{aligned}$$

$$H(Y) = H(\rho)+(1-\rho)\cdot H(\alpha)$$

Casos particulares:

Si $\rho = 0 \Longrightarrow H(Y) = H(\alpha) = H(X)$ canal sin errores

Si $\rho = 1 \Longrightarrow H(Y) = 0$, canal sin capacidad de transmisión, siempre se recibe un borrón

b) $H(X|Y) = \sum_i P(Y=i)\cdot\sum_j P(X=j|Y=i)\cdot\log_2\frac{1}{P(X=j|Y=i)}$,

donde $i \in \{0, B, 1\}$, $j \in \{0, 1\}$

Calculamos primero la entropía condicional por un valor concreto de la salida Y:

$$H(X|Y=i) = \sum_j P(X=j|Y=i)\cdot\log_2\frac{1}{P(X=j|Y=i)}$$

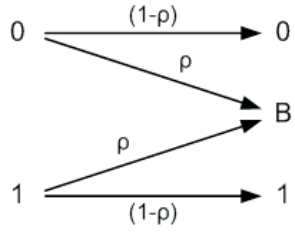


Fig. 1.17: Esquema de transmisión de datos

Para hallar la entropía condicional de cada valor, se determinan las probabilidades de la entrada para un valor de la salida:

$$\left.\begin{array}{l} P(X=0|Y=0)=1 \\ P(X=1|Y=0)=0 \end{array}\right\} \Longrightarrow H(X|Y=0)=0$$

$$\left.\begin{array}{l} P(X=0|Y=B)=\dfrac{p(\alpha)}{p(\alpha)+p(1-\alpha)}=\alpha \\ P(X=1|Y=B)=\dfrac{p(1-\alpha)}{p(\alpha)+p(1-\alpha)}=1-\alpha \end{array}\right\} \Longrightarrow H(X|Y=B)=H(\alpha)$$

$$\left.\begin{array}{l} P(X=0|Y=1)=0 \\ P(X=1|Y=1)=1 \end{array}\right\} \Longrightarrow H(X|Y=1)=0$$

Finalmente, obtenemos:

$$H(X|Y)=P(Y=B)\cdot H(\alpha)=\rho\cdot H(\alpha)$$

c) $$H(Y|X)=\sum_j P(X=j)\cdot\sum_i P(Y=i|X=j)\cdot\log_2\frac{1}{P(Y=i|X=j)},$$

donde $j\in\{0,\,1\},\ i\in\{0,\,B,\,1\}$

Calculamos la entropía condicional de la salida para cada posible valor de entrada:

$$H(Y|X=j)=\sum_i P(Y=i|X=j)\cdot\log_2\frac{1}{P(I=i|X=j)}$$

$$\begin{aligned} H(Y|X=0) &= P(Y=0|X=0)\cdot\log_2\left(\frac{1}{P(Y=0|X=0)}\right)+ \\ &+ P(Y=B|X=0)\cdot\log_2\left(\frac{1}{P(Y=B|X=0)}\right) \end{aligned}$$

$$H(Y|X=0)=(1-\rho)\cdot\log_2\frac{1}{1-\rho}+\rho\cdot\log_2\frac{1}{1-\rho}=H(\rho)$$

$$H(Y|X=1)=H(\rho),\quad \text{por simetría}$$

Luego,

$$H(Y|X)=\alpha\cdot H(\rho)+(1-\alpha)\cdot H(\rho)=H(\rho)$$

Como cabía esperar, la entropía a la salida cuando se conoce la entrada depende únicamente de la probabilidad de borrado del canal:

$$H(Y|X) = H(\rho)$$

d) La información mutua se puede derivar a partir de las entropías condicionales halladas anteriormente:

$$I(X;Y) = H(X) - H(X|Y) = H(\alpha) - \rho \cdot H(\alpha) = (1-\rho) \cdot H(\alpha)$$

o también:

$$I(X;Y) = H(Y) - H(Y|X) = H(\rho) + (1-\rho) \cdot H(\alpha) - H(\rho)$$

$$I(X;Y) = (1-\rho) \cdot H(\alpha)$$

e) La capacidad del canal se obtiene directamente a través de su definición:

$$C = \text{máx}_{\alpha}(1-\rho) \cdot H(\alpha) = (1-\rho) \cdot H_{\text{máx}}$$

$$H_{\text{máx}} = 1, \text{ cuando } \alpha = \frac{1}{2}$$

$$C = (1-\rho) \text{ bits}$$

Problema 10

Un sistema de transmisión de datos utiliza un regenerador de señal. El regenerador tiene por entradas (X) símbolos que pertenecen al alfabeto $\{1, 0, -1\}$. Las probabilidades de recepción de los símbolos son:

$$P[X=1] = \alpha,\ P[X=0] = 1-\alpha-\beta,\ P[X=-1] = \beta$$

para $0 < \alpha + \beta <= 1$.

El regenerador restituye los valores de los borrones ($X = 0$) en valores de salida $Y = 1$ o $Y = -1$, con la misma proporción con la que se generan, y mantiene el mismo valor ($Y = X$) cuando las entradas son $X = 1$ o $X = -1$. Así, el sistema de transmisión de datos regenerador se puede caracterizar a través de la matriz estocástica de probabilidades de transición:

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ \frac{\alpha}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \\ 0 & 1 \end{pmatrix} \quad 0 < \alpha + \beta \leq 1$$

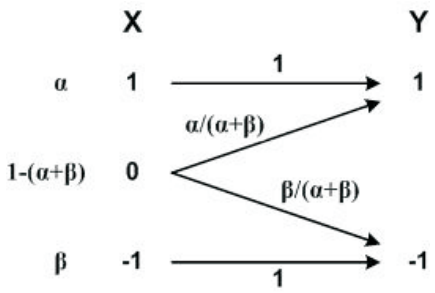


Fig. 1.18: Esquema de transmisión de datos del regenador de símbolos

a) Determine $H(Y)$.
b) Calcule $H(Y|X)$.
c) Halle $I(X;Y)$.
d) Calcule la capacidad del sistema regenerador en bits por símbolo para los casos:
I) $\alpha = \beta$
II) $\alpha = 2 \cdot \beta$

Solución

a) Determinamos $H(Y)$:

$$P(Y=1) = \alpha + [1-(\alpha+\beta)] \cdot \frac{\alpha}{\alpha+\beta} = \alpha + \frac{\alpha}{\alpha+\beta} - \alpha = \frac{\alpha}{\alpha+\beta}$$

$$P(Y=-1) = 1 - P(Y=1) = \frac{\beta}{\alpha+\beta}$$

$$H(Y) = \frac{\alpha}{\alpha+\beta} \cdot \log_2 \frac{\alpha+\beta}{\alpha} + \frac{\beta}{\alpha+\beta} \cdot \log_2 \frac{\alpha+\beta}{\beta} \triangleq H\left(\frac{\alpha}{\alpha+\beta}\right)$$

b) Hallamos $H(Y|X)$:

$$H(Y|X) = P(X=1) \cdot H(Y|X=1) + P(X=-1) \cdot H(Y|X=-1) + P(X=0) \cdot H(Y|X=0)$$

$$H(Y|X) = P(X=0) \cdot H(Y|X=0)$$

dado que los otros sumandos son 0, ya que no hay incertidumbre cuando $X=1$ o $X=-1$. Así:

$$H(Y|X) = [1-(\alpha+\beta)] \cdot H\left[\frac{\alpha}{\alpha+\beta}\right] = [1-(\alpha+\beta)] \cdot H(Y)$$

c) $I(X;Y) = H(Y) - H(Y|X) = H(Y) - [1-(\alpha+\beta)]H(Y) = (\alpha+\beta) \cdot H(Y)$

$$I(X;Y) = (\alpha+\beta) \cdot H\left(\frac{\alpha}{\alpha+\beta}\right) = (\alpha+\beta) \cdot \left[\frac{\alpha}{\alpha+\beta} \cdot \log_2 \frac{\alpha+\beta}{\alpha} + \frac{\beta}{\alpha+\beta} \cdot \log_2 \frac{\alpha+\beta}{\beta}\right]$$

d) ara $\alpha = \beta$

$$I(X;Y) = 2 \cdot \alpha \cdot \left[\frac{1}{2} \cdot \log_2 2 + \frac{1}{2} \cdot \log_2 2\right] = 2 \cdot \alpha, \qquad \alpha + \beta \leq 1 \Rightarrow \alpha \leq 1/2$$

$$C = \text{máx}_\alpha \, I(X;Y), \text{ con } 0 < \alpha \leq \frac{1}{2}$$

$$C = I(X;Y)|_{\alpha = \frac{1}{2}} = 1 \text{ bit}$$

En este caso, el repetidor nunca recibe borrados.

Para $\alpha = 2 \cdot \beta$, no se reciben borrados si $1 - \alpha - \beta = 0$, por lo que $1 - 3 \cdot \beta = 0 \Rightarrow \beta = \frac{1}{3}$ y $\alpha = \frac{2}{3}$

$$I(X;Y)|_{\alpha=2\cdot\beta} = \beta \cdot [3 \cdot \log_2 3 - 2] \Rightarrow C = \frac{1}{3}[3 \cdot \log_2 3 - 2] = 0{,}91 \text{ bits}$$

Problema 11

Se desea analizar el comportamiento de n canales binarios simétricos (BSC), con probabilidad de error p, conectados en serie, como se muestra en la figura 1.19.

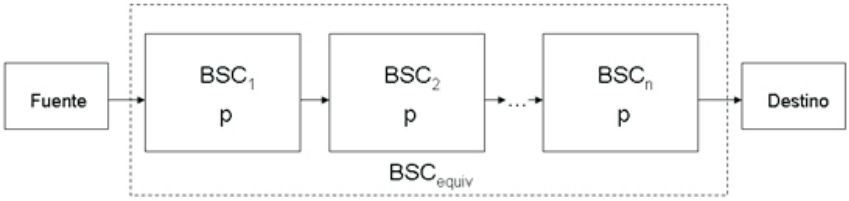


Fig. 1.19: Canales binarios simétricos en serie

Responda a las siguientes preguntas:

a) Para el caso de dos canales BSC en cascada ($n = 2$), determine la matriz estocástica de probabilidades del canal BSC equivalente.

b) Determine la probabilidad de error del BSC equivalente, p_{equiv}, cuando $n = 2$. Halle la capacidad del canal BSC equivalente para este caso, $n = 2$.

c) Para el caso el caso general, en que $p \ll 1/n$, obtenga un valor aproximado de p_{equiv} que dependa solo de n y p. Para este caso, especifique una expresión simple de la capacidad del canal BSC equivalente.

Solución

a) $n = 2$

$$Q_{BSC_1} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \triangleq Q_{n=1}$$

En cascada:

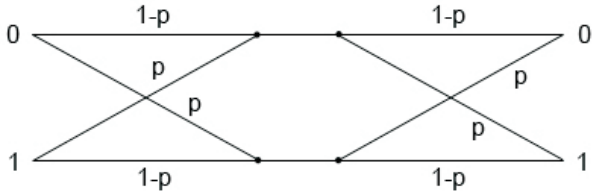


Fig. 1.20: Esquema de transición de datos para dos canales BSC

$$\begin{aligned} p(s=0|e=0) &= (1-p)^2 + p^2 \\ p(s=1|e=0) &= p(1-p) + (1-p)p = 2p(1-p) \\ p(s=1|e=1) &= (1-p)^2 + p^2 \\ p(s=0|e=1) &= p(1-p) + (1-p)p = 2p(1-p) \end{aligned}$$

$$Q_{n=2} = \begin{bmatrix} (1-p)^2 + p^2 & 2p(1-p) \\ 2p(1-p) & (1-p)^2 + p^2 \end{bmatrix} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \cdot \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

Se observa que $Q_{n=2} = Q_{n=1}^2$

En general, $\boxed{Q_n = Q_n^n = 1}$

b) $n = 2$

$$Q_{\text{equiv}} = \begin{bmatrix} 1-p_{\text{equiv}} & p_{\text{equiv}} \\ p_{\text{equiv}} & 1-p_{\text{equiv}} \end{bmatrix} = \begin{bmatrix} (1-p)^2 + p^2 & 2p(1-p) \\ 2p(1-p) & (1-p)^2 + p^2 \end{bmatrix}$$

Identificando $p_{\text{equiv}} = 2p(1-p) = 2p - 2p^2$.

La capacidad para el BSC es $C = 1 - H(p)$.

Para el $\text{BSC}_{\text{equiv}}$ será $C = 1 - H(p_{\text{equiv}}) = 1 - H(2p - 2p^2)$.

c) Si $p \ll 1/n < 1$, entonces $p^i \ll p$, con $i = 2, 3, 4, \ldots$

Considerando que se produce un error en recepción cuando hay un número impar de errores en los n canales, podemos hallar la expresión general. Así:

$$p_i \triangleq prob[i \text{ errores en } n \text{ canales}] = \binom{n}{i} p \frac{1}{2} (1-p)^{n-1}$$

$$p_{\text{equiv}} = p_1 + p_3 + p_5 + \cdots + p_{2k-1} \quad k = \left\lfloor \frac{n+1}{2} \right\rfloor$$

$$p_{\text{equiv}} = \sum_{k=1}^{\left\lfloor \frac{n+1}{2} \right\rfloor} p_{2k-1} = \sum_{k=1}^{\left\lfloor \frac{n+1}{2} \right\rfloor} \binom{n}{2k-1} p^{2k-1} (1-p)^{n-2k+1}$$

Aproximando con $p \gg p^3 \gg p^5 \ldots$ $p_{\text{equiv}} \simeq n\, p\, (1-p)^{n-1}$

Considerando que $1 - p \simeq 1$, entonces $p_{\text{equiv}} \simeq n\, p$ y $C = 1 - H(np)$

Problema 12

Considere tres canales discretos con los siguientes diagramas de transiciones:

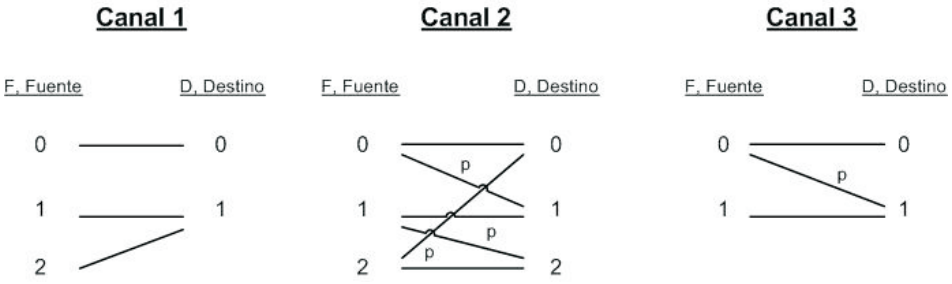


Fig. 1.21: Diagramas de transición

a) Calcule la capacidad de canal para cada canal. Exprésela en función de p para los canales 2 y 3.

b) Calcule las tres capacidades de canal anteriores para $p = 1/2$. ¿Con qué canal se puede transmitir más información por cada uso que se haga de él?

Nota: Para mayor claridad de la solución, llame

$$H(p) = p \cdot \log_2 \frac{1}{p} + (1-p) \cdot \log_2 \frac{1}{1-p}$$

Solución

a) Hallamos las matrices de probabilidad de transición $p(D|F)$ para cada canal:

Canal 1	Canal 2	Canal 3
$P(D\|F) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$	$P(D\|F) = \begin{pmatrix} 1-p & p & 0 \\ 0 & 1-p & p \\ p & 0 & 1-p \end{pmatrix}$	$P(D\|F) = \begin{pmatrix} 1-p & p \\ 0 & 1 \end{pmatrix}$
Canal simétrico respecto a la entrada	Canal simétrico	

Tabla 1.4: Matrices de probabilidades de transición

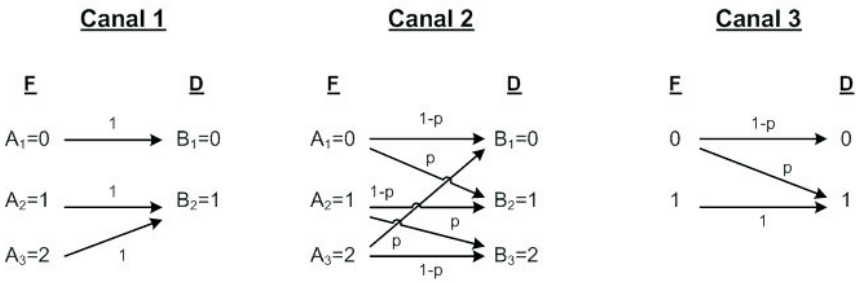


Fig. 1.22: Diagramas de transición

b) Calculamos la información mutua de la fuente a la entrada con respecto a la salida del canal. Esta información mutua o transmisión de información es la disminución de la entropía a la entrada del canal si conocemos la salida. También es la disminución de la entropía a la salida del canal si conocemos la entrada.

$$I(F;D) = H(F) - H(F|D) = H(D) - H(D|F) \text{ [bits/símbolo]}$$

La capacidad de canal se define como la máxima transmisión de información, por lo que

$$C = \mathrm{m\,x}_{p\{A_i\}} I(F;D) = \mathrm{m\,x}_{p\{A_i\}} [H(D) - H(D|F)] \text{ [bits/símbolo]}$$

donde $p\{A_i\}$ es el conjunto de probabilidades de los símbolos emitidos por la fuente F.

Canal 1

$$H(D|F) = \sum_i p(A_i) \cdot H(D|A_i) = \{\overbrace{H(D|A_1) = H(D|A_2) = H(D|A_3)}^{\text{Canal determinista} \Rightarrow H(D|F)=0} 1 \cdot \log_2 1 = 0\}$$

$$= 0 = H(D|A_1)$$

Se trata de un canal determinista, en el que, conocida la entrada, la salida queda determinada:

$$H(D|F) = 0$$

$$H(D) = \sum_j P(B_j) \cdot \log_2 \frac{1}{P(B_j)}$$

Por tanto, para hallar la capacidad de canal hay que maximizar la entropía a la salida, en función de la estadística de la fuente de entrada.

$$C = \mathrm{m\,x}_{p\{A_i\}}[H(D) - H(D|F)] = \mathrm{m\,x}_{p\{A_i\}} H(D)$$

$$p(B=0|A=1) = P(B=0|A=2) = 0$$

$$p(B=0|A=0) = 1$$

$$\begin{aligned} p(B=0) = {} & p(B=0|A=0) \cdot p(A=0) + p(B=0|A=1) \cdot p(A=1) + \\ & + p(B=0|A=2) \cdot p(A=2) = p(A=0) \end{aligned}$$

$$p(B=1|A=0) = 0$$

$$p(B=1|A=1) = p(B=1|A=2) = 1$$

$$\begin{aligned} p(B=1) = {} & p(B=1|A=0) \cdot p(A=0) + p(B=1|A=1) \cdot p(A=1) + \\ & + p(B=1|A=2) \cdot p(A=2) = p(A=1) + p(A=2) \end{aligned}$$

$$C = \mathrm{m\,x}_{p\{A_i\}}[H(D) - H(D|F)] = \mathrm{m\,x}_{p\{A_i\}} H(D)$$

Para hacer máxima la información mutua $I(F;D)$, debemos maximizar la $H(D)$. Podemos probar si existe una F, con $p(A_i)$ coherentes, tal que la fuente sea equiprobable: $p(B=0) = p(B=1)$.

$$\left\{ \begin{array}{c} p(B=0) = p(B=1) = p(A=0) = p(A=1) + p(A=2) \\ p(B=0) + p(B=1) = 1 \end{array} \right\} \Rightarrow$$

$$\Rightarrow p(B=0) = p(B=1) = p(A=0) = 1/2, \quad p(A=1) + p(A=2) = 1/2$$

Será para una F tal que $p\{A_i\}$ tengan los valores anteriores. Por tanto, sí que existe una F tal que $H(D)$ llega al valor máximo posible. En caso contrario, habría que maximizar la expresión resultante.

Entonces, $H(D) = 2 \cdot \frac{1}{2} \log_2 2 = 1$ bit/símbolo

$C = 1$ bit

Canal 2

$$H(D|F) = \sum_i p(A_i) \cdot H(D|A_i)$$

$$H(D|A_1) = H(D|A_2) = H(D|A_3) = p \cdot \log_2 \frac{1}{p} + (1-p) \cdot \log_2 \frac{1}{1-p} = H(p)$$

por lo que

$$H(D|F) = \sum_i p(A_i) \cdot H(D|A_i) = H(p) \cdot \sum_i P(A_i) = H(p)$$

$$H(D) = \sum_j p(B_j) \cdot \log_2 \frac{1}{p(B_j)}$$

$$\begin{cases} p(B=0) = p(A=0) \cdot (1-p) + p(A=2) \cdot p \\ p(B=1) = p(A=0) \cdot p + p(A=1) \cdot (1-p) \\ p(B=2) = p(A=2) \cdot (1-p) + p(A=1) \cdot p \end{cases}$$

$$C = \text{máx} I(F;D) = \text{máx}_{p\{A_i\}}[H(D) - H(D|F)] = \text{máx}_{p\{A_i\}} H(D) - H(p)$$

donde vemos que $H(p)$ es constante e independiente de $P\{A_i\}$.

Para obtener C, hemos de hacer $H(D)$ máxima $\Rightarrow$ Podemos probar si existe una F, $p(A_i)$, tal que haga que D sea equiprobable, $p\{B_j\} = 1/3\ \forall j$:

$$\text{¿}\exists\ F,\ \ p\{A_i\} \mid p\{B_j\} = 1/3?$$

Sí, para $p\{A_i\} = \frac{1}{3}, \forall i \ \Rightarrow\ P\{B_j\} = \frac{1}{3}\,\forall j$, tal como vemos en las ecuaciones

$p(B=0)$, $p(B=1)$ y $p(B=2)$.

Entonces, $H(D) = 3 \cdot \frac{1}{3} \cdot \log_2 3 = \log_2 3 = 1{,}5849$ bits

$C = (1.5849 - H(p))$ [bits/símbolo]

Canal 3

$$H(D|F) = \sum_i p(A_i) \cdot H(D|A_i) = p(A=0) \cdot H(D|A=0) + p(A=1) \cdot H(D|A=1)$$

$$H(D|A=0) = \sum_j p(B_j|A=0) \cdot \log_2 \frac{1}{p(B_j|A=0)} = (1-p) \cdot \log_2 \frac{1}{1-p} +$$

$$+ \quad p \cdot \log_2 \frac{1}{p} = H(p)$$

$$H(D|A=1) = \sum_j p(B_j|A=1) \cdot \log_2 \frac{1}{p(B_j|A=1)} = 0 + 1 \cdot \log_2 1 = 0$$

$$H(D|F) = p(A=0) \cdot H(p) + p(A=1) \cdot 0 = p(A=0) \cdot H(p)$$

Vemos que $H(D|F)$ no es constante, pues depende de $p(A=0)$, es decir, de la estadística de la fuente F.

$$\begin{aligned} H(D) &= \sum_j p(B_j) \cdot \log_2 \frac{1}{p(B_j)} \\ &= \{p(B=0) = (1-p) \cdot p(A=0),\ p(B=1) = p \cdot p(A=0) + p(A=1)\} \\ &= (1-p) \cdot p(A=0) \cdot \log_2 \frac{1}{(1-p) \cdot p(A=0)} + \\ &+ (p \cdot p(A=0) + p(A=1)) \cdot \log_2 \frac{1}{p \cdot p(A=0) + p(A=1)} \end{aligned}$$

$$\begin{aligned} C &= \mathrm{m\,x}_{p\{A_i\}}[H(D) - H(D|F)] = \{\text{sea } z = p(A=0)\} \\ &= \mathrm{m\,x}_{p\{A_i\}}[(1-p) \cdot z \cdot \log_2 \frac{1}{(1-p) \cdot z} + (p \cdot z + (1-z)) \cdot \\ &\cdot \log_2 \frac{1}{p \cdot z + (1-z)} - z \cdot H(p)] \Longrightarrow C(z) \end{aligned}$$

¿Para qué $p\{A_i\}$ la $I(F;D)$ se hace máxima? Hemos de maximizar $C(z)$ y ese valor máximo nos dará C:

$$C(z) = (1-p) \cdot z \cdot \log_2 \frac{1}{(1-p) \cdot z} + (z \cdot p + 1 - z) \cdot \log_2 \frac{1}{(z \cdot p + 1 - z)} - z \cdot H(p)$$

(Ver **nota 2**)

$$\begin{aligned} \text{C'(z)} &= (1-p) \cdot \log_2 \frac{1}{(1-p)z} + \frac{(1-p)^2 \cdot z^2}{\ln 2} \cdot \left(\frac{-1 \cdot (1-p)}{(1-p)^2 z^2} \right) + (p-1) \cdot \\ &\cdot \log_2 \frac{1}{zp + 1 - z} + \frac{(zp + 1 - z)^2}{\ln 2} \cdot \left(\frac{-p+1}{(zp + 1 - z)^2} \right) - H(p) \\ &= (1-p) \log_2 \frac{1}{(1-p) \cdot z} - \frac{1-p}{\ln 2} + (p-1) \cdot \log_2 \frac{1}{zp + 1 - z} + \frac{1-p}{\ln 2} - H(p) \\ &= (1-p) \cdot \left[\log_2 \frac{1}{(1-p) \cdot z} - \log_2 \frac{1}{1 - (1-p) \cdot z} \right] - H(p) \\ &= (1-p) \log_2 \frac{1 - (1-p) \cdot z}{(1-p) \cdot z} - H(p) \end{aligned}$$

$$C'(z) = 0 \longrightarrow \log_2 \frac{1-(1-p)\cdot z}{(1-p)\cdot z} = \frac{H(p)}{1-p} \longrightarrow 2^{\frac{H(p)}{1-p}} = \frac{1-(1-p)\cdot z}{(1-p)\cdot z}$$

$$2^{\frac{H(p)}{1-p}} \cdot (1-p)\cdot z = 1-(1-p)\cdot z$$

$$\left(2^{\frac{H(p)}{1-p}} + 1\right)\cdot (1-p)\cdot z = 1$$

$$z_{máx} = \frac{1}{(1-p)\cdot\left(2^{\frac{H(p)}{1-p}} + 1\right)} = p(A=0)$$

(Ver **nota 1**)

Ya hemos encontrado la fuente F, con una distribución de probabilidades $p\{A_i\}$ que maximiza $I(F;D)$:

$p(A=0) = z_{máx}(p)$

$p(A=1) = 1 - z_{máx}(p)$ Así, la capacidad de canal, $C(Z = Z_{m\,x})$ tiene esta expresión,

que depende del parámetro p.

$$C(p) = (1-p)\cdot z_{máx}\cdot \log_2 \frac{1}{(1-p)z_{máx}} + (p\cdot z_{máx} + (1-z_{máx}))\cdot$$

$$\cdot \log_2 \frac{1}{p\cdot z_{máx} + (1-z_{máx})} - z_{máx}\cdot H(p)\ [\text{bits}]$$

c) $p = 1/2$

Canal $1 \rightarrow C_1 = 1$ bit

Canal $2 \rightarrow C_2 = 1.5849 - H(p = 1/2) = 0{,}5849$ bit

Canal $3 \rightarrow H(p = 1/2) = \frac{1}{2}(\log_2 2)\cdot 2 = 1$

$$z_{máx} = p(A=0) = \frac{1}{\frac{1}{2}(2^{\frac{1}{1/2}} + 1)} = \frac{2}{5} = 0.4;\ p(A=1) = 0.6$$

$$p(B=0) = \frac{1}{2}\cdot 0.4 = 0.2$$

$$p(B=1) = \frac{1}{2}\cdot 0.4 + 0.6 = 0.8$$

$$C_3 = \frac{1}{2}\cdot 0.4\cdot \log_2 \frac{2}{0.4} + \left(\frac{1}{2}\cdot 0.4 + 0.6\right)\cdot \log_2 \frac{1}{0.8} - 0.4$$
$$= 0.2\cdot \log_2 5 + 0.8\cdot \log_2 1.25 - 0.4 = 0{,}3219 \text{ bits/símbolo}$$

Con el canal 1, podemos transmitir más cantidad de información por cada uso que se haga del mismo.

Nota 1: Para ser estrictos, faltaría comprobar que $z_{\text{máx}}$ ofrece un máximo en $F(z)$:

$F''(z_{\text{máx}} < 0)$

Otra manera de verlo es si $F(z) < C$ para un $z < z_{\text{máx}}$ y un $z > z_{\text{máx}}$, y que solo $F(z_{\text{máx}}) = C$.

Como $(z_{\text{máx}} = 0.6)$, $F(0.3) = 0.3008 < C = 0.3219$

Del mismo modo, $F(0.5) = 0.3115 < C = 0.3219$

Nota 2

$$\log_a x = \frac{\ln x}{\ln a}, \quad (\ln x)' = \frac{1}{x}, \quad (\log_a x)' = \left(\frac{\ln x}{\ln a}\right)' = \frac{1}{\ln a}\cdot\frac{1}{x}$$

Problema 13

Una fuente emite dos símbolos (A, B) y queda completamente caracterizada por las siguientes probabilidades de emisión condicionadas:

$p(A|A) = 0.6$
$p(A|B) = 0.3$

Dicha fuente atraviesa un canal binario simétrico, **C**, con una tasa de error de $p = 0.2$.

Se pide:

a) ¿Cuál es la entropía de la fuente?

b) ¿Cuál es la entropía a la salida del canal? Coméntese el resultado. Se decide utilizar tres canales idénticos a C en paralelo según el esquema:

c) ¿Cuál es la capacidad del nuevo canal mostrado en la figura 1.23?

d) ¿Cuál es la entropía a la salida del nuevo canal? Coméntese el resultado.

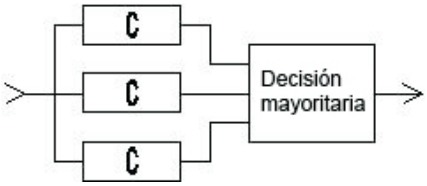


Fig. 1.23: Disposición de los canales

Solución

a) Primero, se calculan las entropías condicionadas:

$$H(X|A) = 0.6 \log_2 \frac{1}{0.6} + 0.4 \log_2 \frac{1}{0,4} = 0{,}971 \text{ bits}$$

$$H(X|B) = 0.7 \log_2 \frac{1}{0.7} + 0.3 \log_2 \frac{1}{0.3} = 0{,}881 \text{ bits}$$

A continuación, se calcula la probabilidad de cada estado de fuente:

$$P(A) = P(A|A)P(A) + P(A|B)P(B)$$

$$P(A) + P(B) = 1$$

$$P(A) = P(A)0.6 + 0.3(1 - P(A)) \Rightarrow P(A) = 0.429,\ P(B) = 0.571$$

Por último, se combinan las entropías condicionadas para obtener el total:

$$H = H(X|A)P(A) + H(X|B)P(B) = 0.971 \cdot 0.429 + 0.881 \cdot 0.571 = 0{,}92 \text{ bits}$$

b) La entropía a la salida se obtendrá combinando las entropías condicionadas:

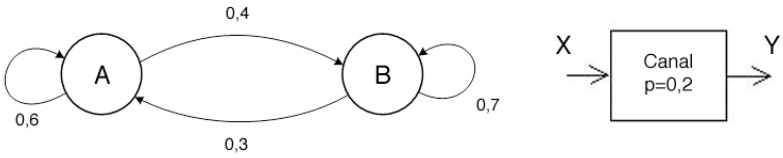


Fig. 1.24: Probabilidades de transición

$$H = H(Y|X = A) \cdot P(A) + H(Y|X = B) \cdot P(B)$$

$$H(Y|X = A) = P(Y = A|X = A) \cdot \log_2 \frac{1}{P(Y = A|X = A)} +$$

$$+ P(Y = B|X = A) \cdot \log_2 \frac{1}{P(Y = B|X = A)}$$

$$H(Y|X=B) = P(Y=A|X=B)\cdot \log_2 \frac{1}{P(Y=A|X=B)} + \\ + P(Y=B|X=B)\cdot \log_2 \frac{1}{P(Y=B|X=B)}$$

$$P(Y=A|X=A) = P(A|A)\cdot(1-p) + P(B|A)\cdot p$$
$$P(Y=B|X=A) = 1 - P(Y=A|X=A)$$
$$P(Y=B|X=B) = P(A|B)\cdot p + P(B|B)\cdot(1-p)$$
$$P(Y=A|X=B) = 1 - P(Y=B|X=B)$$

Resolviendo:

$$P(Y=A|X=A) = 0.56$$
$$P(Y=B|X=A) = 0.44$$
$$P(Y=B|X=B) = 0.62$$
$$P(Y=A|X=B) = 0.38$$

$H(Y|X=A) = 0.990$

$H(Y|X=B) = 0.958$

$H(Y) = 0{,}972$ bits

Puesto que el canal genera incertidumbre, la entropía a su salida es mayor.

c) El nuevo canal se comportará como un canal binario simétrico, con una probabilidad de cruce que deberá ser menor. La probabilidad de cruce del nuevo canal será la probabilidad de que en los canales en paralelo se produzcan dos o tres cruces.

$$p_e = \binom{3}{2}p^2(1-p) + \binom{3}{3}p^3 = 3\cdot p^2(1-p) + 1\cdot p^3 = 3p^2 - 2p^3 = 0.104$$

$$\text{Capacidad del nuevo canal} = 1 - \left[0.104\log_2\frac{1}{0.104} + 0.896\log_2\frac{1}{0.896}\right]$$
$$= 0{,}518 \text{ bits/simb}$$

d) Repitiendo el apartado *b* con $p = 0.104 \Rightarrow H(Y) = 0{,}95$ bits/simb

Al ser un canal mejor, la entropía a su salida es menor.

Problema 14

Un código ternario utiliza las longitudes ($l_1 = 3, l_2 = 2, l_3 = 3, l_4 = 3, l_5 = 3, l_6 = 3$) para unos símbolos con probabilidades de ocurrencia ($p_1 = 1/4, p_2 = 1/6, p_3 = 1/12, p_4 = 1/6, p_5 = 1/4, p_6 = 1/12$), respectivamente. Sin extender la fuente, puede decirse que:

a) La longitud media es inferior a $H + 1$.

b) Cumple la desigualdad de Kraft, por lo que es instantáneo.

c) No existe otro código con longitud media menor.

d) Ninguna de las respuestas anteriores es correcta.

Solución

a) La longitud media de codificación será:

$$\overline{L} = 3 \cdot \frac{1}{4} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{12} + 3 \cdot \frac{1}{6} + 3 \cdot \frac{1}{4} + 3 \cdot \frac{1}{12} = \overline{L} = \frac{2}{6} + 3 \cdot \frac{5}{6}$$

$$= 2{,}833 \text{ dígitos ternarios}$$

$$H + 1 = \sum_{\forall i} p_i \cdot \log_3 \; p_i + 1 = \cdots = 1.551 + 1$$

$$= 2{,}55 \text{ dígitos ternarios/símbolo} \rightarrow a \text{ falsa}$$

b) Que el código cumpla la desigualdad de Kraft no garantiza que este sea instantáneo, sino que existe un código con esas mismas longitudes que lo es. Por tanto *b* es falsa.

c) Por inspección en el árbol

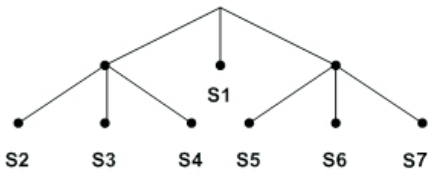


Fig. 1.25: Codificació ternaria de Huffman

$\rightarrow \overline{L} = 2 \cdot \frac{2}{6} + 1 \cdot \frac{2}{6} = 1{,}75$ dígitos ternarios/símbolo. Por tanto, *c* es falsa.

2
Codificación de canal

2.1. Introducción

El contenido teórico correspondiente a este capítulo presenta los principales conceptos, técnicas y algoritmos que se utilizan en la codificación y la descodificación de la información transmitida. En multitud de situaciones, el índice de error del sistema de transmisión sin codificar resulta demasiado elevado. En tal caso, es necesario recurrir a técnicas de codificación de canal, bien para detectar errores (y realizar una retransmisión de los datos), bien para corregirlos [COST83, GITL92].

Los métodos de codificación de canal se basan siempre en la introducción de una cierta redundancia en la secuencia de información, lo que implica una disminución de la tasa de transmisión, a igual esquema modulador. En recepción, el descodificador aprovecha la redundancia para determinar si ha habido errores y, en su caso, intentar corregirlos. Los codificadores pueden dividirse en dos categorías básicas: códigos bloque (sistema combinacional o sin memoria) y códigos convolucionales (sistema secuencial o con memoria) [SKLA88].

Los códigos bloque trabajan con una cantidad fija de símbolos de información y añaden cierta cantidad de símbolos redundantes en función del número de símbolos que pueden corregir o detectar. Los códigos convolucionales pueden interpretarse como un filtro digital. Por consiguiente, los codificadores convolucionales aceptan la secuencia de entrada de forma continuada y generan una salida de tasa mayor [CARL86].

Puede plantearse otra estrategia en cuanto a la protección de la información que se vierte al canal. Así, la redundancia puede añadirse en el propio proceso de modulación, aumenta el número de puntos de la constelación e imbrica los procesos de modulación y codificación. Este tipo de técnicas se conocen básicamente con el nombre de TCM (*trellis-coded modulation* o modulación codificada por enrejado) y poseen la interesante propiedad de no alterar la tasa real de transmisión, a igual potencia y ancho de banda.

2.2. Contenidos teóricos

Este es la relación de los contenidos teóricos que se cubren en clase de transmisión de datos.

- Fundamentos básicos. Estrategia FEC (*forward error correction*) *versus* ARQ (*automatic repeat request*)
- Códigos bloque
  - Capacidad correctora/detectora de errores y correctora de borrados
  - Entrelazado
  - Códigos e-perfectos y de Hamming
  - Códigos polinómicos y códigos cíclicos
- Códigos convolucionales
- Modulación codificada

2.3. Bibliografía

[CARL86] Carlson, Bruce A., *Communication Systems*, McGraw-Hill Int., 4¦ ed., ISBN-10: 0070111278, 2001.

[COST83] Lin, S.; Costello, J., *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, ISBN-10: 013283796X, 1983.

[GITL92] Gitlin, Richard D.; Hayes, Jeremiah F.; Weinstein, Stephen B., *Data Communications Principles*, Plenum Press, Nueva York, Boston, ISBN-10: 0306437775, 1992.

[SKLA88] Sklar, B., *Digital Communications. Fundamentals and Applications*, Prentice Hall, 2¦ ed., ISBN-10: 0130847887, 2001.

2.4. Problemas

Problema 1

Sea un LFSR con el polinomio de conexiones primitivo $c(D) = 0103$ (en notación octal, mayor peso a la izquierda). El estado inicial del LFSR es D^2. ¿Cuánto vale $D^{192} \bmod c(D)$?

a) 1
b) D^3
c) D^5
d) Ninguna de las respuestas anteriores es correcta

Solución

Es habitual expresar los polinomios de conexiones en notación octal. Para hallar el polinomio, basta con expresarlo en binario y asociar coeficientes de $c(D)$ a las potencias correspondientes a los unos.

$c(D) = 0103$. Codificación de 0103 en binario $= 0\,0\,0|0\,0\,1\,|0\,0\,0|0\,1\,1| = D^6 + D + 1$

Por tanto, $S^0(D) = D^2$. Por el enunciado, sabemos que el polinomio $c(D)$ es primitivo, con lo que producirá secuencias periódicas de longitud máxima:

$L = L_{\text{máx}} = 2^m - 1 = 2^6 - 1 = 63$, siendo m el grado del polinomio de conexiones $c(D)$.

Como sabemos que $D^L \cdot p^0(D) \bmod c(D) = p^0(D)$, en particular también $D^L \bmod c(D) = 1$, escribimos:

$D^{192} \bmod c(D) = D^{3\cdot 63+3} \bmod c(D) = D^3 \bmod c(D) = D^3 \bmod D^6 + D + 1 = D^3$

Problema 2

Un bibliotecario está introduciendo los códigos ISBN-10 de varios libros en una aplicación. Al introducir el ISBN del libro *Digital Communications* de E. Lee y D. Messerschmitt, observa que hay un dígito rasgado imposible de leer: $0792 * 93910$. ¿Qué afirmación es cierta?

a) El código ISBN correcto es 0792893910.
b) El valor correcto del borrón es 4.
c) No es posible corregir ese borrón.
d) Ninguna de las respuestas anteriores es correcta.

Solución

El código ISBN tiene capacidad correctora de borrones $\rho = 1$ igual a la capacidad detectora de errores, $\delta = 1$. No corrige ningún error ($e = 0$).

Sea $Z = 0792 * 93910$ la palabra recibida, donde $*$ indica un borrón en la palabra. Como el número de borrones no es superior a la capacidad correctora de borrones del código ($1 \leq \rho = 1$), ese borrón se puede corregir.

Para proceder a su corrección, basta con igualar el síndrome asociado a la palabra recibida a 0:

$\vec{s_r} = Z \cdot H^\top = \vec{0_r}$

donde H es la matriz de comprobación del código ISBN vista en clase y r es la redundancia del código, $r = 1$. Denominamos a la incógnita, es decir, el valor del borrón.

$s = Z \cdot H^T = 0$

$$s = (0792a93910) \cdot \begin{pmatrix} 10 \\ 9 \\ 8 \\ 7 \\ 6 \\ 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} = 235 + 6a = 0 \bmod 11$$

, ya que el código ISBN trabaja en Z_{11}

$Z_{11} \rightarrow \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, X\}$, siendo $X \equiv 10$

Probamos múltiplos de 11 tales que a sea un entero que haga cumplir la ecuación:

$(235 + 6a) \bmod 11 = 0$

¿$235 + 6a = 242$? $\Rightarrow$ $6a = 7$ No es a un entero.

¿$235 + 6a = 253$? $\Longrightarrow$ $6a = 18$ $\Longrightarrow$ $a = 3$

El código ISBN correcto es 0792393910. La respuesta correcta es *d*.

Problema 3

Sea un código de Hamming sistemático con la siguiente matriz de comprobación, tal que $H = (-P^T | I_r)$:

$$H = \begin{pmatrix} 110 * * * * \\ 011 * * * * \\ 101 * * * * \end{pmatrix}$$

Se transmite $Y = 0000000$ y durante la transmisión se producen errores en las posiciones 2, 3, 4 y 5. ¿Qué mensaje de usuario descodificaríamos?

a) $X = 0100$
b) $X = 0111$
c) $X = 0011$
d) $X =$ Ninguno de los anteriores

Solución

Observamos que $H_{r\times n} \longrightarrow r = 3$ filas; $n = 7$ columnas $\Longrightarrow k = n - r = 4$

La matriz de comprobación de un código sistemático cumple la siguiente forma. Recuérdese que la matriz de comprobación $H_{r\times n}$ está formada por r vectores de n componentes, linealmente independientes. Así, H genera el subespacio vectorial ortogonal al código. Se cumple que $G_{k\times n} \cdot H^T_{n\times r} = \varnothing_{k\times r}$ es la matriz nula de k filas y r columnas. Ambas matrices son ortogonales.

Para $H = (-P^T|I_r)$, se cumple:

$$H_{r\times n} = (-P^T|I_r) \Longrightarrow H = \left(\begin{array}{c:c} 110* & 100 \\ 011* & 010 \\ 101* & 001 \end{array}\right) \Longrightarrow \left(\begin{array}{c} * \\ * \\ * \end{array}\right) = \left(\begin{array}{c} 1 \\ 1 \\ 1 \end{array}\right)$$

Vemos que esta es la columna que falta, ya que un código de Hamming ha de tener todas las columnas de H diferentes para poder corregir los n errores simples. Ya conocemos H:

$$\Longrightarrow H = \left(\begin{array}{c:c} 1101 & 100 \\ 0111 & 010 \\ 1011 & 001 \end{array}\right)$$

. Palabra código enviada: $Y = 0\ \underline{0}\ \underline{0}\ \underline{0}\ \underline{0}\ 0\ 0 \Longrightarrow Z = 0\ 1\ 1\ 1\ 1\ 0\ 0$ es la palabra recibida, dado que las posiciones señaladas experimentan errores.

Hallamos el síndrome asociado a Z:

$$\vec{s_r} = Z \cdot H^T = (0\ 1\ 1\ 1\ 1\ 0\ 0) \cdot H^T = \left(\begin{array}{c} 101 \\ 110 \\ 011 \\ 111 \\ 100 \\ 010 \\ 001 \end{array}\right) = (1\ 1\ 0) = 2^{\text{a}} \text{ fila de } H^T \Longrightarrow$$

$$\Longrightarrow \hat{\vec{e}} = 0\ 1\ 0\ 0\ 0\ 0\ 0$$

Vemos que $\vec{s_r}$ coincide con la segunda fila de H^T, por lo que se resuelve que el vector de error debió ser un $\vec{e_n}$ con $n - 1$ ceros y un 1 en la posición del error, el segundo bit.

Finalmente, puesto que la palabra recibida Z es el resultado de sumar el vector de error a la palabra enviada Y, es decir:

$$Z = Y \oplus \vec{e}$$

podemos despejar la palabra que estimamos que fue enviada:

$$\hat{Y} = Z \oplus \hat{\vec{e}} = 0111100 + 0100000 = \underbrace{0011}_{k=4}100$$

El mensaje estimado, al ser un código sistemático, coincide con los k primeros bits de la palabra código estimada:

$$\hat{X} = 0011$$

Problema 4

Sea un código polinómico sistemático (7,4), con polinomio generador $g(D) = 1 + D + D^3$. Halle la matriz generadora.

a) $\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

d) Ninguna de las respuestas anteriores es correcta.

Solución

Se trata de un código cíclico sistemático (7,4) $\longrightarrow$ $r = n - k = 3$, ya que $n = 7$ y $k = 4$

Al ser sistemático, la matriz generadora tiene esta forma:

$$G_{k\times n} = (I_k | P_{k\times r}) = \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & - & - & - \\ 0 & 1 & 0 & 0 & - & - & - \\ 0 & 0 & 1 & 0 & - & - & - \\ 0 & 0 & 0 & 1 & - & - & - \end{array}\right)$$

En un código cíclico sistemático, se cumple:

$R(D) = D^r \cdot X(D) \bmod g(D)$

$Y(D) = D^r \cdot X(D) + R(D)$

siendo $X(D)$ el mensaje de usuario y $C(D)$ el polinomio de conexiones, escritos ambos polinómicamente.

Por tanto, para obtener la matriz generadora $G_{k\times n}$, basta con obtener la codificación de los cuatro mensajes de usuario 1000, 0100, 0010 y 0001, correspondientes a la base canónica del conjunto de mensajes de usuario.

a) $X = 1000 \equiv D^3 = X(D),\ D^r \cdot X(D) = D^6$

$R(D) = D^r \cdot X(D)$ mód $g(D) = D^6$ mód $(D^3 + D + 1) = D^2 + 1$

D^6 | D^3+D+1

$D^6+D^4+D^3$ | D^3+D+1

D^4+D^3

D^4+D^2+D

D^3+D^2+D

D^3+D+1

$D^2+1=R(D)$

$Y(D) = D^r \cdot X(D) + R(D) = D^6 + D^2 + 1 \equiv 1000|101$

b) $X = 0100 \equiv D^2 = X(D),\ D^r \cdot X(D) = D^5$

$R(D) = D^r \cdot X(D)$ mód $g(D) = D^5$ mód $(D^3 + D + 1) = D^2 + D + 1$

$Y(D) = D^r \cdot X(D) + R(D) = D^5 + D^2 + D + 1 \equiv 0100|111$

c) $X = 0010 \equiv D = X(D),\ D^r \cdot X(D) = D^4$

$R(D) = D^r \cdot X(D)$ mód $g(D) = D^4$ mód $(D^3 + D + 1) = D^2 + D$

D^4 | D^3+D+1

D^4+D^2+D | D

$D^2+D=R(D)$

$Y(D) = D^r \cdot X(D) + R(D) = D^4 + D^2 + D \equiv 0010|110$

d) $X = 0001 \equiv 1 = X(D),\ D^r \cdot X(D) = D^3$

$R(D) = D^r \cdot X(D)$ mód $g(D) = D^3$ mód $(D^3 + D + 1) = D + 1$

$$\begin{array}{l|l} D^3 & D^3+D+1 \\ \hline D^3+D+1 & 1 \\ \hline D+1=R(D) & \end{array}$$

$Y(D) = D^3 + D + 1 \equiv 0001|011$

Finalmente, esta es la matriz generadora, por lo que la respuesta correcta es la *a*.

$$G_{k\times n} = (I_k|P_{k\times r}) = \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right)$$

Problema 5

Sea un código (7,4) de Hamming, cuya matriz de comprobación (a la que le falta una columna por determinar, marcada con asteriscos) es:

$$H = \left(\begin{array}{ccccccc} 0 & 1 & * & 1 & 1 & 0 & 0 \\ 1 & 1 & * & 1 & 0 & 1 & 0 \\ 1 & 1 & * & 0 & 0 & 0 & 1 \end{array}\right) = (-P^T|I_r)$$

El demodulador detecta una alta presencia de ruido en dos muestras, que marca como *a* y *b*. En cada caso, se recibe la palabra *Z* y se estima el mensaje de usuario *X*. ¿Qué caso es posible que se haya producido?

a) $Z = 1a0b101,\ X = 1101$
b) $Z = ab11011,\ X = 0111$
c) $Z = 1ab0110,\ X = 1010$
d) Ninguna de las respuestas anteriores es correcta.

Solución

Al ser un código de Hamming (1-perfecto), la capacidad correctora de errores es $e = 1$. El vector de error (errores corregibles) $\vec{e_n}$ tendrá una única componente no nula. Los síndromes asociados a cada error corregible coinciden con las columnas de $H_{r\times n}$, por lo que estas deben ser diferentes. Así, cada error corregible tiene un síndrome asociado diferente.

$$\vec{s_r} = Z \cdot H^\top = (Y + \vec{e}) \cdot H^\top = Y \cdot H^\top + \vec{e} \cdot H^\top = \vec{e_n} \cdot H^\top_{n\times r}$$

Hamming ⇒ e=1 ⇒ H tiene las 7 columnas diferentes.

Por tanto, la columna que falta en $H_{r\times n}$ es la terna de bits que queda disponible:

$$\begin{pmatrix} * \\ * \\ * \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Hamming $\longrightarrow$ 1-perfecto $\longrightarrow \left\{ \begin{array}{l} d_{min} = 2e + 1 = 3 \\ \rho = d_{min} - 1 = 2 \end{array} \right\} \Longrightarrow$ siempre puede corregir hasta dos borrados.

Para corregir los borrados presentes en una palabra recibida Z_n, basta con forzar que el síndrome asociado $\vec{s}_r$ sea nulo:

$\vec{s}_r = Z_n \cdot H^{\top}_{n\times r} \equiv 0$

Una vez corregida Z_n, el mensaje estimado $\hat{X}_k$ se corresponde con los $k = 4$ primeros bits de Z_n.

a) $\vec{s}$=Z·H^T=(1a0b101)$\cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (1 + a + b, 1 + a + b, a) \equiv (000) \quad \Rightarrow$

$\left\{ \begin{array}{l} a = 0 \\ b = 1 \end{array} \right\} \quad \Rightarrow \quad X = 1001$. No coincide.

b) $\vec{s} = Z \cdot H^T = (b,\ a + b,\ a + b) \equiv (000) \quad \Longrightarrow \quad a = b = 0 \quad \Longrightarrow \quad X = 0011.$

No coincide.

c) $\vec{s} = Z \cdot H^T = (1+a+b,\ a,\ 1+a+b) \equiv (000) \quad \Longrightarrow \quad \left\{ \begin{array}{l} a = 0 \\ b = 1 \end{array} \right\} \quad \Rightarrow \quad X = 1010.$

Respuesta correcta.

Problema 6

Un sistema de transmisión utiliza un código corrector de Hamming (7,4). Si la probabilidad de error de bit en el canal es 10^{-4}, ¿cuánto vale la tasa de error de bit a nivel de usuario?

Solución

Los códigos de Hamming son 1-perfectos, es decir, corrigen hasta un error y nunca corrigen más. Por tanto, para que se produzca al menos un error en el bloque, después de descodificar es necesario que se produzcan dos o más errores en el bloque recibido.

$$p_E(\text{bloque}) = \sum_{i=2}^{7} \binom{n}{i} p^i (1-p)^{7-i}.$$

Esta expresión, puesto que el valor de p es suficientemente bajo, se puede aproximar por

$$p_E(\text{bloque}) \approx \binom{7}{2} p^2(1-p)^5 = 21 \cdot (10^{-4})^2(1-10^{-4})^5 \approx 21 \cdot 10^{-8}$$

Ahora bien, cuando se producen dos errores en el bloque, el codificador de Hamming lo interpreta como un error simple, siempre en una posición distinta de donde se hallaban los dos errores, ya que el síndrome que obtiene es la suma de las dos columnas de la matriz de comprobación en las que se produjeron los errores. Como la suma de cualesquiera columnas de dicha matriz siempre es una tercera columna, el descodificador introducirá un nuevo error, por lo que aparecerán un total de tres errores en el bloque descodificado.

Por todo ello, la probabilidad de error de bit será:

$$\begin{aligned} p_e(\text{bit}) &= \frac{\#\text{bits erróneos}}{\#\text{bits totales}} = \frac{3 \cdot \#\text{bloques erróneos}}{7 \cdot \#\text{bloques totales}} \\ &= \frac{3}{7} \cdot 21 \cdot 10^{-8} = 9 \cdot 10^{-8} = \frac{3}{7} \cdot 21 \cdot 10^{-8} = 9 \cdot 10^{-8} \end{aligned}$$

Problema 7

En un código de Hamming (7,4) sistemático, puede afirmarse que:

a) La submatriz de paridad puede tener dos filas iguales.
b) La submatriz de paridad puede tener dos columnas iguales.
c) La matriz de comprobación puede tener dos filas iguales.
d) Nada de lo anterior puede afirmarse.

Solución

Al ser de Hamming, todas las columnas de H han de ser diferentes, puesto que es 1-perfecto:

$$H(3 \times 7) = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} = (-P^T | I_r)$$

$$G(7\times 4)=\begin{pmatrix}1&0&0&0&1&1&0\\0&1&0&0&0&1&1\\0&0&1&0&1&1&1\\0&0&0&1&1&0&1\end{pmatrix}=(I_k|P)$$

donde la segunda submatriz de G es la submatriz de paridad P. Por lo tanto las filas de P sólo pueden ser (1 0 1), (1 1 1), (1 1 0) i (0 1 1) en cualquier orden.

a) 2 filas de P iguales $\Longrightarrow$ 2 columnas de $\left\{\begin{array}{c}P^T\\H\end{array}\right\}$ iguales $\Longrightarrow$ imposible.

b) 2 columnas de P iguales $\Longrightarrow$ las 4 filas de P son 4 vectores particulares y no se puede conseguir reordenando las filas o las columnas.

c) Imposible, pues está I_r en $H \Longrightarrow$ 2 filas de H no son iguales nunca.

Problema 8

El polinomio de conexiones de un LFSR es $D^4 + D + 1$. Indique la respuesta *falsa*:

a) La secuencia generada es $D^{11} + D^8 + D^7 + D^5 + D^3 + D^2 + D + 1$.
b) La probabilidad de que exista un 0 es de 7/15.
c) La secuencia generada tiene ráfagas de cuatro 1 y tres 0.
d) Alguna de las respuestas anteriores es falsa.

Solución

Puesto que D^4+D+1 es primitivo, genera una secuencia de máximo período (MLSR), en este caso igual a $2^4 - 1 = 15$. Además, en este tipo de secuencias, la probabilidad de emitir un 0 es $p(0) = 7/15$ y la de emitir un 1 es $p(1) = 8/15$. Por tanto b es cierta.

La secuencia de salida puede calcularse con la siguiente operación:

$$\begin{array}{l|l} D^{15}+1 & D^4+D+1 \\ \hline 0 & D^{11}+D^8+D^7+D^5+D^3+D^2+D+1 \Longrightarrow \text{a) cierta} \end{array}$$

por lo que a es cierta.

Del resultado anterior, se obtiene que la secuencia de salida es 000100110101111, es decir c es cierta.

Por todo lo explicado, se concluye que d es *falsa*.

Problema 9

Sea un código (n, k) que se caracteriza porque la distancia entre dos palabras cualesquiera es cuatro. Se puede afirmar que:

a) El código es 2-perfecto.
b) El código es 4-perfecto.
c) El código es 1-perfecto.
d) Nada de lo anterior es cierto.

Solución

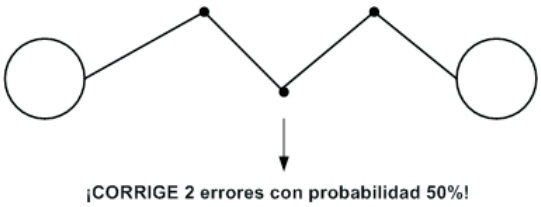


Fig. 2.1: Esquema de corrección para código e-perfecto

Un código es e-perfecto cuando corrige hasta e errores y nunca $e + 1$. Si la distancia del código es par, cuando se reciba una n-pla a distancia $\frac{d_{\text{mín}}}{2} > e = \left\lfloor \frac{d_{\text{mín}-1}}{2} \right\rfloor$, se corregirá en el 50 % de los casos, por lo que ningún código con distancia par puede ser perfecto. Se concluye, por tanto, que la respuesta correcta es la d.

Problema 10

Un código de Hamming (7,4) se ha extendido con 1 bit de paridad global para utilizarlo en un canal con probabilidad de error de bit de 10^{-4} y una probabilidad de borrón de 10^{-3}. La probabilidad p de recibir un error y un borrón es:

a) $p \geq 0.044 \cdot 10^{-3}$
b) $0.044 \cdot 10^{-3} > p \geq 0.033 \cdot 10^{-3}$
c) $0.033 \cdot 10^{-3} > p \geq 0.022 \cdot 10^{-3}$
d) $0.022 \cdot 10^{-3} > p$

Solución

Para realizar este cálculo, es necesario encontrar un número de palabras con un borrón y un error, esto es, $\binom{8}{2}$ multiplicado por 2 (no es lo mismo un borrón en la posición i y un error en la j que un borrón en la posición j y un error en la i).

Cualquiera de estos casos en los que se produce un borrón y un error tiene una probabilidad de que ocurra igual a $p_e \cdot p_b\,(1-(p_e+p_b))^6$. Es decir:

$$2\cdot\binom{8}{2}p_e\cdot p_b\,(1-(p_e+p_b))^6 = 2\cdot\binom{8}{2}p_e^2\cdot(1-2p_e)^6$$
$$= \{\text{con } p_b = p_e = 10^{-3}\} = 0.055\cdot 10^{-3}$$

Por tanto, la solución correcta es *a*.

Problema 11

Se dispone de un código (6,3) binario lineal y sistemático, corrector de errores, y sean Y_1, Y_2 palabras código, donde $Y_1 = 110110$ y $Y_2 = 101011$. Calcúlese:

a) Capacidad correctora del código. ¿Es un código perfecto?
b) ¿Cómo se codifica el mensaje 111?
c) Indíquese si la siguiente matriz puede ser de comprobación:

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

d) Supóngase que se ha recibido la palabra $Z = 001001$. ¿Cuál sería la decisión del descodificador si se utiliza el código como corrector?

Solución

Puesto que se trata de un código corrector de errores $\Rightarrow \left\{ \begin{array}{c} e \geq 1 \\ d_{min} \geq 3 \end{array} \right\} \Rightarrow W_{min} = 3$

Disponemos de dos palabras código $Y_1 = 110110$ y $Y_2 = 101011$, pero el código es (6, 3) lineal $\Rightarrow$ necesitaríamos otra más, independiente de las anteriores, para disponer de una base (ya que $k = 3$). Vamos a intentar ver las posibilidades de las ocho palabras código ($k = 3$). Para ello, tendremos en cuenta que el código es sistemático, lineal y corrector de errores:

$\underline{X}$		$\underline{Y}$
000		000 000
001		001 abc $\rightarrow$ Y_3
010		010 $\overline{a}b\overline{c}$ $\rightarrow$ $Y_1 + Y_2 + Y_3$
011	$\longrightarrow$	011 101 $\rightarrow$ $Y_1 + Y_2$
100		100 $a\overline{b}\overline{c}$ $\rightarrow$ $Y_1 + Y_3$
101		101 011 $\rightarrow$ Y_1
110		110 110 $\rightarrow$ Y_2
111		111 $\overline{a}\overline{b}c$ $\rightarrow$ $Y_2 + Y_3$

¿a, b, c? La siguiente tabla muestra todas las combinaciones posibles. Posteriormente, iremos eliminando aquellas que no satisfagan que $d_{\text{mín}} \geq 3$:

a	b	c
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

$Y_3 \rightarrow$ del conjunto (a, b, c), al menos dos han de ser iguales a 1.

$Y_1 + Y_2 + Y_3 \rightarrow$ Del conjunto $(\overline{a}, b, \overline{c})$ al menos dos han de ser iguales a 1.

Casos posibles:

Primera fila: No, Y_3 tendría $W_H = 1$
Segunda fila: No, Y_3 tendría $W_H = 2$
Tercera fila: No, Y_3 tendría $W_H = 2$
Cuarta fila: No, $Y_1 + Y_3$ tendría $W_H = 1$
Quinta fila: No, Y_3 tendría $W_H = 2$
Sexta fila: No, $Y_1 + Y_2 + Y_3$ tendría $W_H = 1$
Séptima fila: Sí
Octava fila: No, $Y_1 + Y_2 + Y_3$ tendría $W_H = 2$

$(a, b, c) = (1, 1, 0)$

$d_{\text{mín}} = 3 \Rightarrow e_c = 1$

Podemos generar todo el código:

000	000	2
001	110	3
010	011	2
011	101	2
100	101	3
101	011	2
110	110	6
111	000	2

a) Puesto que $d_{\text{mín}} = 3 \rightarrow e = \dfrac{d_{\text{mín}} - 1}{2} = 1$ ($e = 1$ ya que cogemos el entero inferior del resultado que obtenemos)

Se puede ver fácilmente que no es un código 1-perfecto. Al ser binario y $e = 1$, para ser perfecto debería cumplir que $2^r = 1 + n$ y, en este caso, $r = 3$, $n = 6$. En realidad, este código es un código recortado obtenido a partir del código(7,4) de Hamming, que sí es 1-perfecto.

b) $X = 111 \rightarrow Y = X \cdot G$

$$G_{kxn} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$Y = (111) \cdot \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} = 111000$$

c) Debe cumplirse que $G \cdot H^T = 0$

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

d) $Z = 001001$

$\vec{s} = Z \cdot H^T = 001001 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 001 \rightarrow$ No coincide con ninguna fila de H^T, por lo que no se confunde con un error simple.

De hecho, vemos que Z está a distancia 2 de varias Y, pero no podemos elegir una Y a distancia mínima. En esta Z, ha habido más errores que la capacidad correctora del código, $e = 1$. El código no confunde ese error múltiple con un error simple, por lo que no la puede corregir. Mejor así, pues si lo hiciera se equivocaría e introduciría otro error. El código sí detecta que esa Z recibida es errónea, pues su síndrome conocido $\vec{s}$ no es nulo.

3
Criptografía

3.1. Introducción

La proliferación del correo electrónico o de los servicios web ha supuesto un cambio sustancial en la forma de difundir la información, que combina información multimedia con enlaces que facilitan el salto a otra página u objeto. La funcionalidad correcta de estos servicios exige una adecuada implantación de medidas de seguridad. La implantación sistemática de servicios de seguridad en las redes existentes requiere la utilización de protocolos y técnicas de seguridad adecuadas, compatibles con las especificaciones presentes. Los servicios de seguridad básicos en las comunicaciones son: autenticación, control de acceso, confidencialidad, integridad de los datos y no repudio [PAS08].

La criptografía es un mecanismo fundamental para implementar los servicios de seguridad antes mencionados. La criptografía, conocida desde antiguo como el arte de la escritura secreta, se ha convertido actualmente en una compa¤era imprescindible del desarrollo de la sociedad de la información. Los objetivos principales a los que sirve la criptografía son la confidencialidad, la integridad y la autenticidad en el tratamiento de la información en formato electrónico. Una de las aplicaciones más notables de esta disciplina es el comercio electrónico seguro [SCH96].

En la figura 3.1 se presenta un esquema de la transmisión segura de un mensaje M entre dos entidades, a través de un canal inseguro.

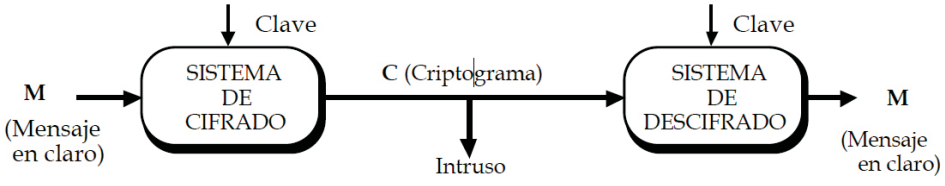


Fig. 3.1: Esquema de transmisión segura de un mensaje

Los sistemas criptográficos de este esquema son los encargados de calcular el mensaje cifrado C, a partir del mensaje en claro M y de la "clave de cifrado", y de realizar el proceso inverso, el descifrado, y así determinar M a partir del mensaje cifrado y la "clave de descifrado". Estas dos claves, como ya veremos más adelante, no tienen que ser necesariamente iguales. Cuando un sistema criptográfico utiliza en el descifrado la misma clave que en el cifrado, se dice que utiliza un "cifrado simétrico". Por el contrario, si la clave de descifrado es distinta a la clave de cifrado, el sistema estará empleando un "cifrado asimétrico" [RIF91].

Un sistema de criptografía simétrico es una familia de transformaciones invertibles, donde emisor y receptor utilizan la misma clave *K*. La clave *K* se ha tenido que poner previamente en conocimiento de las dos partes mediante el uso de un canal secreto. Esta clave necesita, pues, ser distribuida con antelación a la comunicación. El coste y el retardo, impuestos por esta necesidad, son los principales obstáculos para la utilización de la criptografía de clave secreta en grandes redes [STA00].

Entre los algoritmos simétricos, podemos destacar los de cifrado en bloque y los de cifrado de flujo. Estos últimos son los más indicados para entornos de alta velocidad de transmisión. Los algoritmos simétricos de cifrado en bloque son los más utilizados en redes de datos, y se pueden clasificar entre "de dominio público"(se publica el algoritmo con todo detalle) o "propietario"(se mantiene en secreto el algoritmo). En el ambiente académico son preferibles los algoritmos de dominio público que se han sometido a un escrutinio intenso por parte de la comunidad criptográfica y no se han visto comprometidos (diferentes algoritmos propietarios de uso comercial se han vulnerado por métodos de ingeniería inversa).

Probablemente el algoritmo criptográfico más utilizado es el AES(*Avanced Encryption Standar*, 2001) que es de dominio público y sustituto del más antiguo DES (*data encryption standard*, 1977). Este último, aún se sigue utilizando en algunos entornos.

En un sistema de *cifrado asimétrico*, también conocido como *de clave pública*, el cifrador utiliza una clave *P*, mientras que el descifrador utiliza una clave distinta *S*. La clave *P* es pública, y la clave *S* es privada e incalculable a partir de *P* en un tiempo prudente (aunque muchos sistemas de este tipo son vulnerables ante el uso eventual de un computador cuántico). El sistema asimétrico posibilita la comunicación en un sentido; para realizar la comunicación en sentido contrario, se necesita otro par de claves secreta-pública. La principal característica que hace interesantes estos métodos frente a los sistemas criptográficos simétricos es que no se precisa el intercambio de secretos entre los dos comunicantes. Los algoritmos de clave pública se basan en la teoría de números y de cuerpos finitos. Gracias a este fundamento matemático, es posible demostrar la seguridad computacional de estos métodos. Uno de los algoritmos asimétricos más utilizados es el RSA (*Rivest-Shamir-Adleman*).

Aunque su uso está muy generalizado, resulta vulnerable delante de la computación cuántica, por lo que se están desarrollando sistemas basados en la *Post-quantum cryptography*

3.2. Contenidos teóricos

- Introducción: seguridad computacional vs. seguridad incondicional
- Servicios de seguridad: privacidad, autenticidad, verificabilidad
- Clave simétrica o secreta
  - Criptografía clásica
  - Cifrado en bloque
  - Cifrado en flujo
- Clave pública o asimétrica
  - Conceptos básicos
  - Diffie-Hellman
  - RSA
- Funciones de hash
- Firma digital
- Autenticación

3.3. Bibliografía

[PAS98] Pastor, J.; Sarasa, M.A., *Criptografía digital. Fundamentos y aplicaciones*, Prensas Universitarias de Zaragoza, 1998.

[RIF91] Rifá, J.; Huguet, Ll., *Comunicación digital*, Masson, 1991.

[SCH96] Schneier, B., *Applied Cryptography*, 2¦ ed., John Wiley & Sons, 1996.

[STA00] Stallings, W., *Network Security Essentials*, Prentice Hall, 2000.

3.4. Problemas

Problema 1

Se dispone de un cifrador de cuatro bits de entrada y cuatro bits de salida que, para una cierta clave k, tiene la siguiente relación de entrada y salida $[M, Ek(M)]$

M	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
$C = E_k(M)$	7	4	1	E	B	8	5	2	F	C	9	6	3	0	D	A

Tabla 3.1: Tabla 3.1: Relacion entrada salida del cifrador

Se pide:

a) ¿Cuál es el tamaño mínimo de la clave para que el cifrado pueda suponerse perfectamente aleatorio?

b) El cifrado del mensaje **FFF** es **6A6**. Razone por qué puede asegurarse que el cifrado no se está utilizando en modo nativo o ECB.

c) Cuando se realiza un encadenado, como en este caso, es usual utilizar un vector de inicialización. Indique qué alternativas utilizaría para este vector inicial y qué ventajas aportan.

d) Sabiendo que las únicas operaciones usadas son $E_k(\cdot)$ y XOR, encuentre de forma razonada las ecuaciones del cifrador y del descifrador. ¿Cuánto vale el vector inicial?

Como función de hash de un mensaje de **n** bloques, se usa el algoritmo:

$$h_i = E_k(M_i + h_{i-1}) \quad i = 1 \ldots n \quad h_0 = 0 \quad H = h_n$$

e) Calcule el hash de mensaje **FFF**. ¿Cuántos mensajes de tres bloques generarán el mismo hash que **FFF** y diferirán únicamente en los dos primeros bloques del mensaje? (M_1 y M_2 distintos de F)

f) Obtenga de forma razonada, y no por pruebas exhaustivas, el valor de M que hace que el mensaje **M1F** tenga el mismo hash que **FFF**.

Solución

a) Para que un cifrador pueda considerarse perfectamente aleatorio, debe existir al menos una clave para cada biyección posible. De esta forma, y puesto que el número de biyecciones posibles es $16! = 20.922.789.888.000$, el tamaño mínimo en bits de la clave ha de ser: N. bits $\log_2 16! = 44.25 \implies$ longitud = 45 bits

b) Porque el cifrado de un mensaje uniforme produce un criptograma no uniforme.

c) Se puede utilizar un número de secuencia o un estampado de tiempos. La ventaja que aportan es que no se producen mensajes estereotipados, es decir, que el mismo texto claro genera criptogramas distintos en cada ocasión.

d) Posibilidades de encadenado:

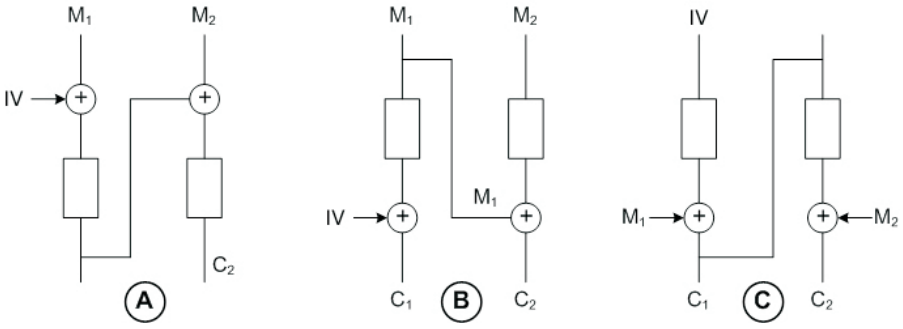


Fig. 3.2: Modos de encadenado

Para cada una de estas posibilidades, se pueden plantear las siguiente ecuaciones:

$$\left.\begin{array}{l} M_1 = F \\ M_2 = F \\ C_1 = 6 \\ C_2 = A \end{array}\right\} \begin{array}{ll} A) & E_k[C_1 + M_2] = E_k[6 + F] = E[9] = C \neq C_2 = A \implies NO \\ B) & E_k[M_2] + M_1 = E_k[F] + F = 5 \neq C_2 = A \implies NO \\ C) & E_k[C_1] + M_2 = E[6] + F = 5 + F = A = C_2 = A \implies OK \end{array}$$

y, por tanto, las ecuaciones del cifrador y del descifrador son, respectivamente:

$$\begin{aligned} C_i &= M_i + E_k(C_{i-1}) \\ M_i &= C_i + E_k(C_{i-1}) \end{aligned}$$

Para encontrar IV, del apartado C, de la figura se deduce la siguiente ecuación:

$$IV = D_k\,[C_1 + M_1] = D_k\,[6 + F] = D_k[9] = A$$

e) Utilizando la ecuación del cálculo del hash:

$$\begin{aligned} h_1 &= E_k[F] = A \\ h_2 &= E_k[F + A] = E_k[5] = 8 \\ h_3 &= E_k[F + 8] = E_k[7] = 2 = H \end{aligned}$$

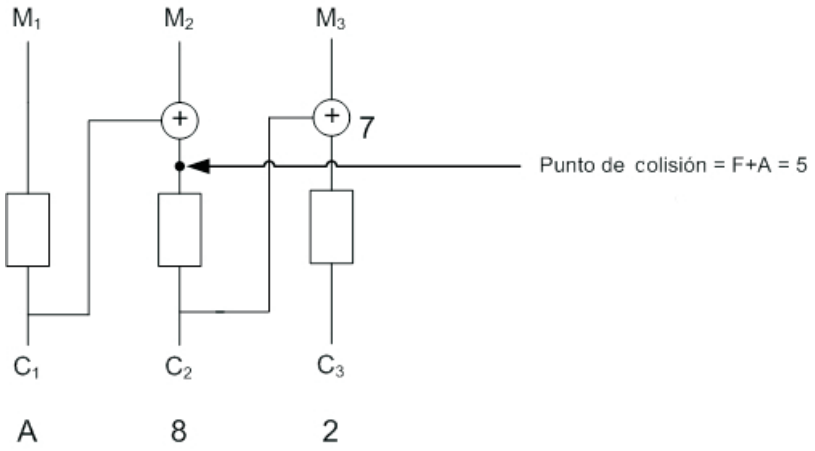


Fig. 3.3: Colisión en las funciones de hash

De la figura 3.3 anterior, se deduce que todos los valores de M_1 y M_2 que satisfagan que en el punto de colisión se obtenga un 5 darán como hash el valor F.

Por tanto, dejando libre M_1(distinto de F), siempre se puede encontrar un valor $M_2(\neq \text{F})$ que genere $H = 2$, por lo que el número de posibles valores para M_1 es 15 (ya que, si M_1 =F, entonces M_2 también ha de valer F.

f) $E_k[M] + 1 = 5 \implies E_k[M] = 4 \implies M = D_k[4] = 1$

Problema 2

Diséñese un registro de desplazamiento de longitud 4 y que genere una secuencia de período máximo.

Solución

Se trata de encontrar un polinomio primitivo de grado 4. Dicho polinomio indicará las conexiones a realizar en el registro de desplazamiento para obtener una secuencia de longitud máxima.

Recuérdese que una condición necesaria para que un polinomio sea primitivo es que sea irreducible. Por ello, comenzaremos por encontrar los polinomios irreducibles de grados 1, 2 y 3 para luego obtener los de grado 4, y de entre ellos buscar uno primitivo. Recuérdese, así mismo, que para cualquier grado existe al menos un polinomio primitivo.

Grado 1

De grado 1, sólo hay dos polinomios posibles, a saber:

1 D
2 $D + 1$

Grado 2

De grado 2, tendremos cuatro posibles polinomios:

1 D^2
2 $D^2 + 1$
3 $D^2 + D$
4 $D^2 + D + 1$

Está claro que el polinomio 1 es divisible por D; por tanto, no puede ser irreducible. Lo mismo sucede para cualquier otro polinomio que no tenga termino independiente. De esta forma, de ahora en adelante podremos eliminar la mitad de los candidatos (de esta lista, eliminamos por la misma razón el polinomio 3).

Veamos qué sucede con el polinomio 2. Por un lado, cualquier polinomio que tenga un número par de términos satisfará que $p(1) = 0$, ya que $1^x = 1$ para cualquier valor de x. Dicho de otra forma, todo polinomio que tenga un número par de coeficientes tendrá 1 como raíz y, por tanto, será divisible por $D + 1$. Para $D^2 + 1$, al tratarse de un cuerpo de característica 2, ocurre que $D^2 + 1 = (D + 1)^2$: la característica 2 hace que, al elevar al cuadrado un polinomio, no aparezcan dobles productos, es decir, que el cuadrado de la suma sea igual a la suma de cuadrados. Por ejemplo:

$$(D + 1)^2 = D^2 + 2D + 1 = D^2 + 1$$

por lo que cualquier polinomio que sólo tenga potencias pares será un cuadrado perfecto y, por tanto, no será irreducible. Veamos un ejemplo:

$$(D^2 + D + 1)^2 = D^4 + D^2 + 1$$

Como resumen, podemos decir que ningún polinomio que cumpla alguna de estas características será irreducible.

De todo lo anterior, se deduce que ninguno de los polinomios 1, 2 y 3 puede ser primitivo, pues son reducibles. Por tanto, el polinomio 4 ha de ser irreducible y primitivo (recuérdese que siempre existe al menos un polinomio primitivo para cualquier grado).

Grado 3

De grado 3, tendremos ocho posibles polinomios:

1	D^3	5	$D^3 + D^2$
2	$D^3 + 1$	6	$D^3 + D^2 + 1$
3	$D^3 + D$	7	$D^3 + D^2 + 1$
4	$D^3 + D + 1$	8	$D^3 + D^2 + D + 1$

Los polinomios 1, 3, 5 y 7 no son irreducibles, pues son divisibles por D. Los polinomios 2, 3, 5 y 8 tampoco lo son porque tienen un número par de términos y, en consecuencia, son divisibles por $D + 1$. Por tanto, nos quedan como únicos candidatos a polinomios irreducibles:

4 $D^3 + D + 1$
6 $D^3 + D^2 + 1$

Ambos deben ser primitivos, porque si un polinomio es primitivo su recíproco también lo es. Se define el recíproco de un polinomio $C(D)$ de grado n y se denota por $C^*(D)$ aquel que satisface que $C^*(D) = D^n C\left(\frac{1}{D}\right)$. Si uno es primitivo, el otro también ha de serlo y, por consiguiente, ambos han de serlo, pues al menos uno de ellos ha de serlo.

Grado 4

De grado 4, tendremos dieciséis posibles polinomios:

1	D^4	9	$D^4 + D^3$
2	$D^4 + 1$	10	$D^4 + D^3 + 1$
3	$D^4 + D$	11	$D^4 + D^3 + D$
4	$D^4 + D + 1$	12	$D^4 + D^3 + D + 1$
5	$D^4 + D^2$	13	$D^4 + D^3 + D^2$
6	$D^4 + D^2 + 1$	14	$D^4 + D^3 + D^2 + 1$
7	$D^4 + D^2 + D$	15	$D^4 + D^3 + D^2 + D$
8	$D^4 + D^2 + D + 1$	16	$D^4 + D^3 + D^2 + D + 1$

La propiedad “no tener término independiente” elimina los polinomios 1, 3, 5, 7, 9, 11, 13 y 15, y los polinomios 2, 3, 5, 8, 9, 12, 14 y 15 son eliminados por la propiedad “tener número par de términos”. En este caso, la propiedad “tener solo potencias pares” elimina los polinomios 2, 5 y 6. Nos quedan como candidatos a irreducibles:

4	$D^4 + D + 1$
10	$D^4 + D^3 + 1$
16	$D^4 + D^3 + D^2 + D + 1$

De estos, vemos que los polinomios 4 y 10 son recíprocos y que el 16 es autorrecíproco. Elegimos el polinomio 4 (también podríamos elegir el 10) porque tiene menos coefi-

cientes. Probaremos si es divisible por los polinomios irreducibles de grado menor que 4 que hemos ido encontrando, es decir:

$$D^2 + D + 1$$
$$D^3 + D + 1$$
$$D^3 + D^2 + 1$$

Comenzando con la prueba, tenemos que:

$(D^4 + D + 1) \mod (D^2 + D + 1) = 1$

$$\begin{array}{l|l} D^4+D+1 & D^2+D+1 \\ \hline D^4+D^3+D^2 & D^2+D \\ \hline D^3+D^2+D+1 & \\ D^3+D^2+D & \\ \hline 1 & \end{array}$$

$(D^4 + D + 1) \mod (D^3 + D + 1) = D^2 + 1$

$$\begin{array}{l|l} D^4+D+1 & D^3+D+1 \\ \hline D^4+D^2+D & D \\ \hline D^2+1 & \end{array}$$

$(D^4 + D + 1) \mod (D^3 + D^2 + 1) = D^2$

$$\begin{array}{l|l} D^4+D+1 & D^3+D+1 \\ \hline D^4+D^3+D & D+1 \\ \hline D^3+1 & \\ D^3+D^2+1 & \\ \hline D^2 & \end{array}$$

Consecuentemente, al no ser divisible por ninguno de los polinomios irreducibles de grado menor, el polinomio $D^4 + D + 1$ es necesariamente irreducible.

A modo de resumen, tenemos los siguientes resultados:

$D^4 + D + 1$	$\mod(D^2 + D + 1) = 1$ $\mod(D^3 + D + 1) = D^2 + 1$ $\mod(D^3 + D^2 + 1) = D^2$
$D^4 + D^3 + 1$	$\mod(D^2 + D + 1) = D$ $\mod(D^3 + D + 1) = D^2$ $\mod(D^3 + D^2 + 1) = D + 1$
$D^4 + D^3 + D^2 + D + 1$	$\mod(D^2 + D + 1) = D + 1$ $\mod(D^3 + D + 1) = D$ $\mod(D^3 + D^2 + 1) = D^2 + 1$

Como se ve, los tres polinomios anteriores son irreducibles. Probaremos si $D^4 + D + 1$ es primitivo. Si no lo fuera, tampoco lo sería su recíproco ($D^4 + D^3 + 1$) y, por tanto, $D^4 + D^3 + D^2 + D + 1$ debería ser forzosamente primitivo.

Para que $D^4 + D + 1$ sea primitivo no ha de dividir a ningún polinomio de la forma $D^\lambda + 1$ para $\lambda < 2^n - 1 = 15$. Si $\lambda < 4$, está claro que ningún polinomio de la forma $D^\lambda + 1$ podrá ser múltiplo de $D^4 + D + 1$.

λ	$D^\lambda+1$	mod(D^4+D+1)
4	D^4+1	D
5	D^5+1	D^2+D+1
6	D^6+1	D^3+D^2+1
7	D^7+1	D^3+D
8	D^8+1	D^2
9	D^9+1	D^3+D+1
10	$D^{10}+1$	D^2+D
11	$D^{11}+1$	D^3+D^2+D+1
12	$D^{12}+1$	D^3+D^2+D
13	$D^{13}+1$	D^3+D^2
14	$D^{14}+1$	D^3
15	$D^{15}+1$	0

Tabla 3.2: Verificación del polinomio $D^4 + D + 1$

Esto demuestra que el polinomio $D^4 + D + 1$ es primitivo. De hecho, no habría sido necesario calcular toda la lista anterior, puesto que los valores pares de λ hacen que $D^\lambda + 1$ sea el cuadrado de $D^{\frac{\lambda}{2}} + 1$ y, por tanto, si éste no era múltiplo, su cuadrado tampoco podría serlo. Así, no era preciso comprobar los valores de λ igual a 4, 6, 8 10, 12 y 14.

Si hubiésemos ensayado con el polinomio $D^4+D^3+D^2+D+1$, hábriamos comprobado que divide D^5+1. Esto es fácil de comprobar:

$(D^5+1) \mod (D^4+D^3+D^2+D+1) = 0$

$$
\begin{array}{l|l}
D^5+1 & D^4+D^3+D^2+D+1 \\
\cline{2-2}
D^5+D^4+D^3+D^2+D+1 & D+1 \\
\cline{1-1}
\quad D^4+D^3+D^2+D+1 & \\
\quad D^4+D^3+D^2+D+1 & \\
\cline{1-1}
\quad 0 &
\end{array}
$$

De aquí se deduce que ningún polinomio que esté completo (tenga todas las potencias) puede ser primitivo, pues si tiene grado n dividirá el polinomio $D^{n+1}+1$.

Problema 3

a) Sea un sistema RSA con los siguientes parámetros para un usuario A ($p = 47$, $q = 59$, $d = 157$). Para la codificación de los mensajes de texto, sustituimos cada letra por un número de dos dígitos, según la siguiente codificación: espacio=00, A = 01, B = 02, ..., Z = 26, y codificamos dos letras por bloque. Codifique el mensaje M = 920 para transmitirlo confidencialmente al usuario A.

b) Realice un cifrado de Vigenére del mensaje M = HOLA con la clave obtenida al descifrar el criptograma generado en el apartado anterior con la clave privada de A.

c) El conocimiento de la función de Euler $\phi(n)$ permite factorizar n en un sistema RSA. Factorice $n = p \cdot q = 2782799$, sabiendo que $\phi(n) = 2779440$. AYUDA: Obtenga $p+q$ como una cierta función de n y $\phi(n)$. Utilice la identidad $(p-q)^2 = (p+q)^2 - 4 \cdot p \cdot q$.

Solución

a) $p = 47; q = 59 \;\; n = p \cdot q = 47 \cdot 59 = 2.773$

$d = 157, \phi(2.773) = 46 \cdot 58 = 2.668$

$e = 17$

blank = 00

A = 01, B = 02, ..., I = 09, T = 20, M = 920,

$C = M^{17} \mod n = 920^{17} \mod 2.773 = 948$

NOTA: Este ejemplo numérico fue presentado en 1977 por Rivest, Shamir y Adleman en su artículo original en el que presentaron el algoritmo RSA

b) $c_i = n_i \oplus k_{i \bmod d} \bmod 27$

En este caso, la longitud de la clave es $d = 2$.

C=QHUU

H	O	L	A	
I	T	I	T	
8	15	12	1	
9	20	9	20	
17	35	21	21	mod 27

0	1	2	3	4	5					10					15					20					25	
_	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

c) $N = 1.879 \cdot 1.481 = 2.782.799$

$\phi(n) = 1.878 \cdot 1.480 = 2.779.440$

Factoriza N.

$p + q = n - \phi(n) + 1 = 3.360$

$p - q = \sqrt{(p+q)^2 - 4 \cdot p \cdot q} = \sqrt{(3.660)^2 - 4 \cdot 2.782.799} = 398$

$2 \cdot p = (p+q) + (p-q) = 3.758 \Rightarrow p = 1.879, \quad q = 1.481$

Esto también aparecía en el artículo de Rivest, Shamir y Adleman.

Problema 4

Sea $H(M)$, con una salida de k bits, que se calcula de la siguiente forma:

1) Se añade al final del mensaje el número de ceros necesarios para que la longitud del mensaje sea múltiplo de k.
2) Se divide el mensaje en n bloques de k bits, m_i $0 \leq i \leq n-1$.
3) $H(M)$ se calcula iterativamente de la siguiente manera:

$$h_0 = m_0$$
$$h_{i+1} = h_i \oplus m_{i+1} \quad 0 \leq i \leq n-2 \text{ (XOR bit a bit)}$$
$$H(M) = h_{n-1}$$

a) Indique las propiedades que debe cumplir una función de hash criptográficamente robusta, y diga cuáles de ellas cumple la función propuesta.

b) Sea el mensaje M = 10101010101010101010. Calcule $H(M)$ para $k = 6$.

c) Sea un sistema de RSA en el que todos los usuarios usan $e = 23$. Genere un par de claves RSA con $p = 11$, $q = 13$. Indique cuáles serían la clave privada y la pública.

d) Firme digitalmente el mensaje del apartado *b* con el sistema de claves generado en el apartado *c* y la función de hash propuesta (considere siempre que los bits de menor peso son los de la derecha). Indique qué servicios de seguridad ofrece la firma digital.

e) Suponga que es un atacante que quiere modificar un mensaje firmado digitalmente con el sistema anterior. Indique la forma más eficiente de hacerlo y genere un mensaje que genere la misma firma que M.

Solución

a)

PROPIEDADES	CUMPLIMIENTO
Entrada cualquier longitud	Sí
Salida longitud fija	Sí
Dado m, es fácil de calcular $H(m)$	Sí
Dado $H(m)$, no podemos encontrar un m que lo genere	NO
No es posible encontrar dos m que generen la misma $H(m)$	NO

b)

$$\begin{aligned} M &= 1\ 0\ 1\ 0\ 1\ 0 \longrightarrow m_0 \\ &= 1\ 0\ 1\ 0\ 1\ 0 \longrightarrow m_1 \\ &= 1\ 0\ 1\ 0\ 1\ 0 \longrightarrow m_2 \\ &= 1\ 0\ 0\ 0\ 0\ 0 \longrightarrow m_3 \end{aligned}$$

$H(m) = 0\ 0\ 1\ 0\ 1\ 0 = 10$ (en decimal) $= m_0 \oplus m_1 \oplus m_2 \oplus m_3$

Nota: Con la función de hash definida, es muy fácil encontrar mensajes que den una función dada. Véase, por ejemplo, el apartado *e* del presente ejercicio.

c) Se trata de un ejemplo clásico de generación de claves RSA:

$N = p \cdot q = 143$; P_A : $e_A = 23$, N_A=143 clave pública

$ed = 1 \mod \phi(N)$ {algoritmo de Euler extendido $\phi(N) = 120$} $\Longrightarrow d = 47$

S_A : $d = 47$ clave privada

d) $M \| E_{SA}(H(m)) = M \| 10^{47} \mod 143 : M \| 43$

FIRMA $= E_{SA}(H(m)) = M^d \mod 143$

Mensaje firmado: 10101010101010101010101011

La firma digital ofrece autenticación, integridad y soporte para no repudio.

e) La forma más eficiente es generar un mensaje que genere la misma $H(m)$.

Por ejemplo: m' =0 0 1 0 1 0

m' puede suplantar a m.

Existen muchas otras posibilidades para generar fácilmente otro m'' tal que $H(m'') = H(m)$. Por ejemplo, añadiendo al mensaje dos bloques de k bits iguales, o bloques de ceros.

Problema 5

Un sistema de votación desde terminales móviles emplea el algoritmo RSA para proporcionar el servicio de verificabilidad a la aplicación. En este sistema, cada terminal móvil dispone de una clave RSA secreta K_s que se emplea para firmar la concatenación del mensaje m y el resumen r. La concatenación es un valor v de 7 bits que se obtiene con la unión de los cuatro bits del mensaje y los tres bits del resumen, de mayor a menor peso ($v = 0\,x\ m_3\ m_2\ m_1\ m_0\ r_2\ r_1\ r_0$). Para determinar el valor del resumen r, se emplea un LFSR con estado inicial nulo y polinomio de conexiones $1 + D + D^3$, el cual se alimenta con los bits del mensaje, empezando con el de mayor peso. Una vez se ha operado en el LFSR con todos los bits del mensaje, el resumen se deriva directamente del polinomio de estado del LFSR, como se muestra en la figura.

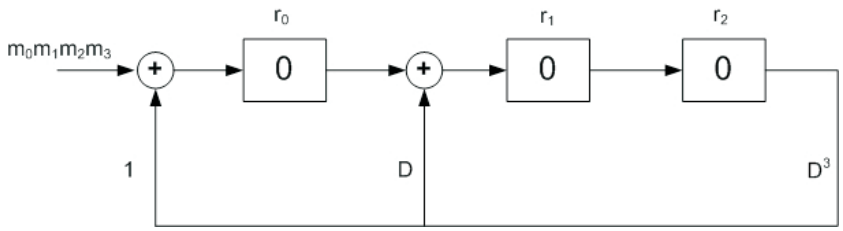


Fig. 3.4: Generación de una función resumen mediante LFSR

Teniendo en cuenta que en un terminal:

$K_s = (d, n) = (7,\ 221) \quad m = 14\ (0x1110)$

a) Calcule la clave pública $K_p = (e,\ n)$ asociada al terminal.

b) Halle el valor del resumen en binario.

c) Especifique el valor concatenado v en decimal. Determine el resultado de la firma de v.

d) Indique cuántos bits son necesarios para enviar cualquier posible valor de la firma de v. Razone la respuesta.

e) A partir de la expresión polinómica para el cálculo iterativo del estado de un LFSR, y con un polinomio $M(D)$ de grado $n-1$ como alimentación externa, obtenga la relación del polinomio de estado en la iteración n, $P^n(D)$, con su valor inicial $P^0(D)$ y con el polinomio $M(D)$.

f) Particularice la expresión anterior para el caso en que el estado inicial del LFSR sea nulo y el valor de $M(D)$ sea $D^7 + D^6 + D^5 + D^4 + D^2 + 1$.

Solución

$k_s = (7, 221), \quad n = 13 \cdot 7 = p \cdot q = 221 \quad m = 14, \quad d = 7, \quad e = ?$

a) $K_p = (e, n)$

$d \cdot e = 1 \mod \phi(n)$

$\phi(n) = (p-1)(q-1) = 192$

$d \cdot e = 1 + K \cdot \phi(n)$

Se obtiene que $e = 55$ y $k = 2$ verifican la ecuación $\Longrightarrow K_p = (55, 221)$

b) Cálculo del resumen

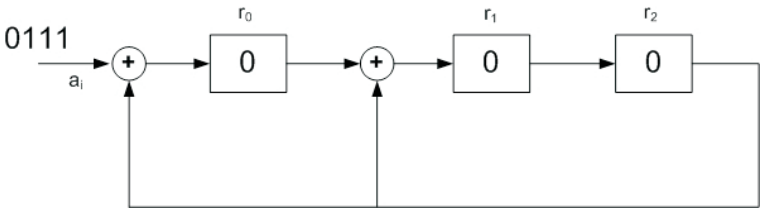


Fig. 3.5: Generación de la función resumen

r =0x101

i	a_i	r_0	r_1	r_2	$P^i(D)$
0		0	0	0	0
1	1	1	0	0	1
2	1	1	1	0	D+1
3	1	1	1	1	D^2+D+1
4	0	1	0	1	D^2+D

Tabla 3.3: Cálculo de la función resumen

c) v =0x1110101=117 $f = v^d \mod n = 117^7 \mod 221 = 195$

d) Puesto que f se reduce modularmente por n, se necesitan tantos bits como sea necesario para codificar el valor $n-1$. En este caso, $n-1=220$, que requiere ocho bits:

$f = 195 = \text{C3h} = \text{0x11000011}$

e) Suponiendo $M(D) = m_{n-1} \cdot D^{n-1} + m_{n-2} \cdot D^{n-2} + \cdots + m_0$

Expresión polinómica para el cálculo iterativo del LFSR:

$P^n(D) = D \cdot P^{n-1}(D) \mod C(D)$

En nuestro caso, cada iteración i da como resultado:

$P^1(D) = D \cdot P^0(D) \mod C(D) + m_{n-1} = (D \cdot P^0(D) + m_{n-1}) \mod C(D)$

$P^2(D) = D \cdot P^1(D) \mod C(D) + m_{n-2} = (D \cdot (D \cdot P^0(D) + m_{n-1}) + m_{n-2}) \mod C(D)$

$P^2(D) = (D^2 \cdot P^0(D) + m_{n-1} \cdot D + m_{n-2}) \mod C(D)$

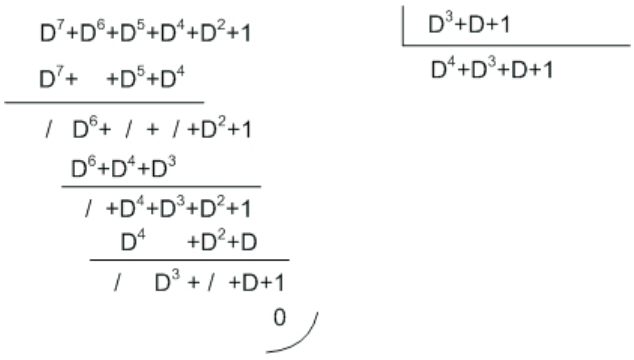
Tras n iteraciones, se obtendrá:

$P^n(D) = (D^n P^0(D) + m_{n-1} D^{n-1} + m_{n-2} D^{n-2} + \cdots + m_0) \mod C(D)$

Agrupando:

$P^n(D) = (D^n P^0(D) + M(D)) \mod C(D)$

d) $P^0(D) = 0$
$M(D) = D^7 + D^6 + D^5 + D^4 + D^2 + 1$
$P^8(D) = M(D) \mod C(D)$



$M(D)$ es múltiplo de $C(D)$.

Problema 6

Sea un sistema de RSA en el que la clave pública del usuario B vale ($N = 7.663$, $e = 4.831$). Utilice la tabla adjunta cuando lo considere necesario:

a) Calcule $X = 397^{1982} \mod 991$. Justifique cómo ha realizado el cálculo.

b) Descodifique el criptograma $C = 000000000101$, enviado por el usuario A al usuario B.

c) Cifre el mensaje $M = 222$ con la secuencia generada por un LFSR caracterizado por el polinomio primitivo $C(D) = D^7 + D + 1$. La clave de sesión determina el estado inicial del LFSR (en este caso, $S(D) = D+1$). Indique posibles debilidades de este cifrador en flujo síncrono, así como la longitud máxima de mensaje que podría cifrarse con una misma clave de sesión.

2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59
61	67	71	73	79	83	89	97	101	103	107	109	113	127	131	137	139
149	151	157	163	167	173	179	181	191	193	197	199	211	223	227	229	233
239	241	251	257	263	269	271	277	281	283	293	307	311	313	317	331	337
347	349	353	359	367	373	379	383	389	397	401	409	419	421	431	433	439
443	449	457	461	463	467	479	487	491	499	503	509	521	523	541	547	557
563	569	571	577	587	593	599	601	607	613	617	619	631	641	643	647	653
659	661	673	677	683	691	701	709	719	727	733	739	743	751	757	761	769
773	787	797	809	811	821	823	827	829	839	853	857	859	863	877	881	883
887	907	911	919	929	937	941	947	953	967	971	977	983	991	997		

Tabla 3.4: Lista de los números primos menores de 1.000

Solución

a) En la tabla adjunta, se observa que 991 es un número primo. Por tanto, $\phi(991) = 990$. Aplicamos la propiedad:

$M^{k\cdot\phi(N)} \mod N = 1 \quad \text{si } mcd(M, N) = 1$

Calculamos $1.982 = 2 \cdot 990 + 2$

$X = 397^{1982}$

$991 \text{ primo} \Longrightarrow \phi(991) = 990$

$\text{mcd}\,(397, 991){=}1 \Longrightarrow 397^{\phi(991)} \mod 991 = 1$

$397^{k\cdot 990} \mod 991 = 1$

$397^{1982} \mod 991 = 397^{2\cdot 990} \cdot 397^2 \mod 991 = 397^2 \mod 991 = 40 = X$

Nótese que el cálculo directo mediante el método del campesino ruso[1] resulta muy tedioso, por lo que es mejor aplicar las propiedades de la función de Euler.

b) Utilizamos la tabla de primos para factorizar N (probando). Posteriormente, tenemos un problema clásico de RSA.

[1] Algoritmo de multiplicación por duplicación que tan solo requiere sumar y hacer mitades. Se basa en descomponer números en potencias de dos y aprovechar que la multiplicación es distributiva, de tal manera que una operación de resultado muy grande se transforma en varias operaciones de resultado menor.

$N = 7663 = 79 \cdot 97$ (tabla de primos), $\phi(N) = 78 \cdot 96 = 7488$

$e = 4831 \implies d = e^{-1}$ mód $\phi(N) = 31$ (algoritmo de Euclides extendido)

$M = C^d$ mód $N = 5^{31}$ mod 7.663=7.476 (en binario): M=1110100110100

c) Criptograma=salida $\oplus$ mensaje

Salida:	0	0	0	0	0	1	1	0
Mensaje:	1	1	0	1	1	1	1	0
Criptograma:	1	1	0	1	1	0	0	0

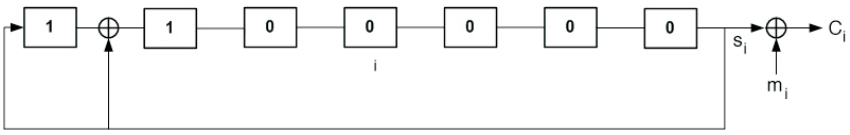


Fig. 3.6: esquema del cifrador en flujo

El cifrador en flujo está formado únicamente por una estructura lineal. Por ello, el criptoanálisis no debe ser complicado. Por otra parte, al ser el polinomio primitivo:

$L_{\text{máx}} = 2^N - 1 = 2^7 - 1 = 127$

Longitud máxima: $L_{\text{máx}} = 127$

Longitud del texto < Long. máx

Problema 7

En un sistema simple de clave pública RSA, se emplea una entidad de certificación (EC) para obtener las claves públicas de las entidades que intervienen en él. Este sistema utiliza en todas las claves públicas el mismo valor $e = 11$, por lo que las claves se reducen a un único valor n. Se ha averiguado que en este sistema todas las claves públicas disponen de un mismo factor primo y que la función resumen empleada es una reducción modular en un cuerpo conmutativo. Sabiendo que la clave pública de la EC es $K_{P_{EC}} = 9.263$ y que un certificado de una entidad A tiene por valor $K_{P_A}|F(R[K_{P_A}]) = 5.959|4.811$:

a) Halle la clave secreta ($K_{S_{EC}}$) de la EC.

b) Calcule el resumen de una clave que incluya el factor primo 127.

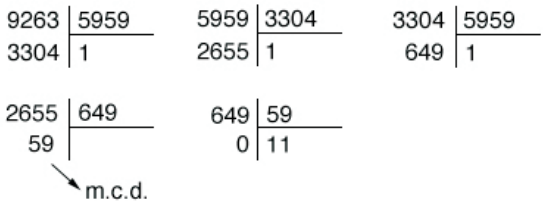
c) Halle la clave pública de valor mínimo en este sistema que tenga la misma firma que la clave anterior. Razone la validez de la función resumen empleada.

Solución

Se determina el factor primo común:

$$\begin{aligned} K_{P_{EC}} &= 9.263 = p \cdot q \\ K_{P_A} &= 5.959 = p \cdot q' \end{aligned} \qquad \text{m.c.d.}(K_{P_{EC}}, K_{P_A}) = p$$

Algoritmo de Euclides:



luego $p = 59$ $\begin{cases} K_{P_{EC}} = 9.263 = 59 \cdot 151 \\ K_{P_A} = 5.954 = 59 \cdot 101 \end{cases}$

a) Derivamos la $K_{P_{EC}}$:

$e \cdot d = 1 + k\,\Phi(K_{P_{EC}})$

$\Phi(K_{P_{EC}}) = 58 \cdot 156 = 9.048$

$11d = 1 + k\,9.048$

$K_1\,9.048 + K_2\,11 = 1$

Algoritmo de Euclides extendido:

$9.048 \cdot 1 + 11 \cdot 0 = 9.048$

$9048 \cdot 0 + 11 \cdot 1 = 11 \quad (-822)$

$9.048 \cdot 1 + 11 \cdot (-822) = 6 \quad (-1)$

$9.048 \cdot (-1) + 11 \cdot (1 + 822) = 5 \quad (-1)$

$9.048 \cdot (1 + 1) + 11 \cdot (-822 - 823) = 1$

$9.048 \cdot 2 + 11 \cdot (-1645) = 1$

$K_2 = -1.645 \Rightarrow d = K_2 \mod 9.048 = 7.403$

$K_{S_{EC}}(d, n) = (7.403, 9.263)$

b) Operación modular realizada:

$r = K_{S_A} \mod m$

Firma obtenida:

$f = E_{K_{S_{EC}}}(r) = 5.811 \Rightarrow r = D_{K_{S_{EC}}}(f)$

$r = 5.811^e \mod n = 4.811^{11} \mod 9.263 = 7$

Se debe verificar: $5.959 \mod m = 7$

donde m es un primo al trabajar en un cuerpo conmutativo.

De forma equivalente:

$5.959 = 7 + km$ $5.952 = km$ Hallamos los factores primos de 5.952 e identificamos:

$$5.952 = 2^6 \cdot 3 \cdot 31 = k\,m \Rightarrow \begin{cases} m = 31 \\ k = 2^6 \cdot 3 \end{cases}$$

$m = 31$ porque debe ser primo y $m > 7$. Por tanto, la operación resumen es:

$r = K_p \bmod 31$

Para una clave $K_p = 127 \cdot 59 = 7.493$ el resumen será, $r = 7.493 \bmod 31 = 22$

c) La K_p mínima que verifica:

$r = K_p \bmod 31 = 22$ se obtiene probando factores primos q pequeños. Con $q = 3$, tenemos:

$K_p 59 \cdot 3 = 177$ y $r = K_p \bmod 31 = 22$

Otra manera: $(59 \cdot q) \bmod 31 = 22 \Rightarrow 59q + 31k = ?$

$59 \cdot 10 - 31 \cdot 19$	=	1	Euclides extendido
$59 \cdot 220 - 31 \cdot 418$	=	22	Multiplicado por 22
$59 \cdot 31 - 31 \cdot 59$	=	0	Ecuación trivial

Combinando las dos últimas ecuaciones:

$59(220 - K_1\,31) - 31(418 - K_1\,49) = 22$

Se busca el valor primo más pequeño de q que cumpla:

$220 - K_1 31 = q \Rightarrow q = 3$ con $K_1 = 7$

Problema 8

Se desea realizar una comunicación de una entidad A a otra entidad B, de forma que los servicios de seguridad diseñados garanticen la integridad del mensaje, la autoría del mensaje y la confidencialidad de la transmisión. Para proporcionar estos servicios las entidades A y B disponen cada una de una clave secreta (K_{SA} y K_{SB}) y una clave pública (K_{PA} y K_{PB}) correspondientes al algoritmo RSA.

La información generada por la entidad A es un bloque de siete bits cuyo valor es 1110110b (76h). La integridad de esta información se garantiza con una función resumen de cinco bits cuyo resultado para el valor de la información mencionado es 01111b (Fh). La firma digital se realiza con cinco bits a partir del valor obtenido en la función resumen.

El mensaje sobre el que se debe garantizar la confidencialidad en el canal de comunicaciones dispondrá de doce bits. Los cinco bits de mayor peso se corresponden con los cinco bits resultantes de la firma digital y los siete de menor peso con los siete de la información.

a) Teniendo en cuenta: $K_{PA} = (e, n) = (17, 33)$; $K_{SA} = (d, n) = (13, 33)$

I) Determine el valor de la firma digital.

II) Exprese en hexadecimal y en decimal el mensaje de doce bits compuesto por A.

III) Demuestre que la elección realizada de las claves K_{PA} y K_{SA} permite que la firma digital sea de solo cinco bits.

b) Sabiendo que los par metros elegidos por la entidad B para calcular su clave pública K_{PB} y su clave secreta K_{SB} son:

$p = 59,\ q = 83,\ e = 11$

I) Razone por qué *e* tiene un valor adecuado, teniendo en cuenta los valores de *p* y *q* elegidos.

II) Halle la clave secreta $K_{SB} = (d, n)$.

III) Calcule cuál es el criptograma enviado por la entidad A a la entidad B. Exprese su valor en hexadecimal.

IV) Comente cuál es el número de bits que se debe asignar a una criptograma en este sistema de acuerdo con las claves elegidas.

Solución

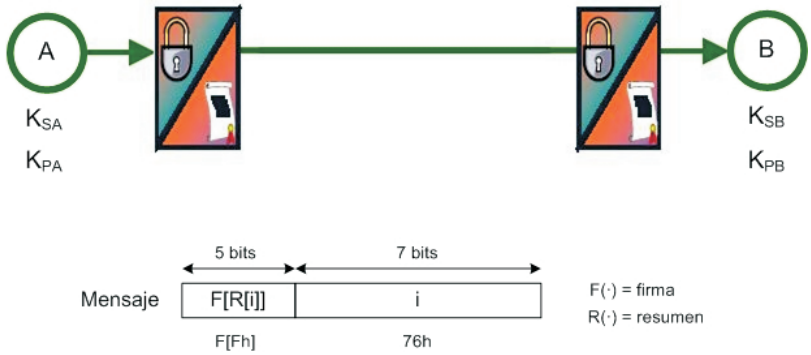


Fig. 3.7: Esquema criptográfico y generación de función resumen (*hash*)

a) $K_{PA} = (e, n) = (17, 33)$, $K_{SA} = (d, n) = (13, 33)$

I) $F(Fh) = 15^{13} \mod 33$

$13 = 1.101_2$

$15^{13} = 15^{2^3+2^2+0\cdot 2^1+1} = \left((15^2 \cdot 15)^2\right)^2 \cdot 15$

$F(Fh) = 9 = 1.001_2$

II) $M = 10011110110_2$

$M = 4F6h$

$M = 1.270$

III) Si $n = 33$, las firmas podrán tener hasta seis bits, ya que según el RSA:

$c = m^e \mod n < n$

Para que los cifrados de valores de cinco bits requieran solo cinco bits es necesario que los cifrados de valores de seis bits den lugar a valores de seis bits. Así,

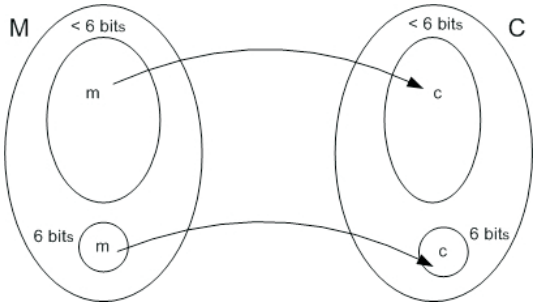


Fig. 3.8: Función resumen

En nuestro caso el único mensaje de seis bits es $m = 32$. Luego, si se verifica:

$32 = 32^e \mod n$ o $32 = 32^d \mod n$, esta hipótesis será cierta.

Nuestro caso cumple que un cifrado de cinco o menos bits da lugar a un resultado de cinco o menos bits.

En este caso: $32^{13} \mod 32 = \left((32^2 \cdot 32)^2\right)^2 \cdot 32 \mod 33 = 32$

b) K_{PB}, K_{SB} a partir de $p = 54$, $q = 83$, $e = 11$

I) m.c.d.$(e, \phi(n)) = 1$; por tanto, se verifica que e y $\phi(n)$ son coprimos:

$\phi(n) = (p - 1) \cdot (q - 1) = 4.756 = 2^2 \cdot 29 \cdot 41$

$e = 11$ m.c.d.$(\phi(n), e) = 1$

II) $K_{PB} = (e, n) = (11, 4.897)$

$K_{SB} = (d, n) = (d, 4.897)$

En RSA, se debe verificar

$e \cdot d = 1 + k \cdot \phi(n) \implies d = e^{-1}$ en $\mathbb{Z}_{\phi(n)}$

Utilizando el algoritmo de Euclides extendido:

$K_1 \cdot \phi(n) + K_2 \cdot e = 1 \implies K_2 = d = 3.459$

III) $C = m^e \mod n$, $K_{PB} = (e, n) = (11, 4.897)$

$C = 1.270^{11} \mod 4.897 = 4.104 = 1.008h$

IV) Para codificar n, necesitamos trece bits:

$n = 4.897 = 1.321h$

Puesto que hay valores de menos de trece bits que dan lugar a criptogramas de trece bits, debemos asignar trece bits para el envío del criptograma.

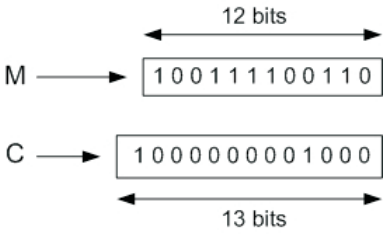


Fig. 3.9: Número de bits para asignar criptograma

Problema 9

Un sistema de firmas digitales utiliza RSA y, como función resumen, el algoritmo denominado El Gamal. Este algoritmo mantiene un valor x en secreto, que debe ser custodiado de igual forma que la clave secreta K_s^{RSA} por la entidad firmante. La verificación de la firma de un mensaje m se lleva a cabo utilizando la clave pública K_p^{RSA} junto con una terna (g, y, p) que facilita la comprobación del mensaje recibido en concordancia con el resumen. En este sistema, será necesario que se hagan públicas las claves K_p^{RSA} y las ternas (g, y, p) asociadas a cada entidad firmante. Considere que el resumen r se concatena a continuación del mensaje m de la forma $m|r$.

Complete el cálculo y la validación del resumen obtenido con el algoritmo El Gamal que se expone, con los siguientes pasos:

1. Se determina un número primo $p = 23$ y dos números aleatorios $g = 15$ y $x = 2$.
2. Se deriva un valor y de la siguiente forma: $y = g^x \mod p$

   a) Determine el valor de y.
3. Para hallar el resumen r de un mensaje $m = 6$, se genera un número aleatorio, coprimo con $p-1$, de valor $z = 3$. A partir de este número, se deriva una primera parte del resumen, denominada a, mediante la expresión: $a = g^z \mod p$

   b) Calcule el valor de a.
4. Se determina un valor auxiliar b', que es elemento inverso de z en el anillo Z_{p-1}

   c) Halle el valor de b'.
5. Se completa el cálculo del resumen con un valor b en Z_{p-1}, que verifica:

   $m = (x \cdot a + z \cdot b) \mod (p - 1)$

   d) Halle el valor de b.
6. Se forma el resumen con la concatenación de los dos valores anteriores, $r = a|b$.
7. La comprobación de un mensaje m se lleva a cabo en el receptor con el resumen r asociado, verificando la igualdad: $y^a \cdot a^b = g^m \mod p$

   e) Compruebe que los cálculos anteriores han sido correctos, utilizando el mecanismo de comprobación del algoritmo.

   f) escriba gráficamente el procedimiento de firma realizado por el emisor y por el receptor.

   g) Razone brevemente la validez de la función resumen propuesta.

Solución

– Se genera un número primo, $p = 23$.

- Se hallan dos números aleatorios, $g = 15$ y $x = 2$.
- Se deriva: $y = g^x \mod p$

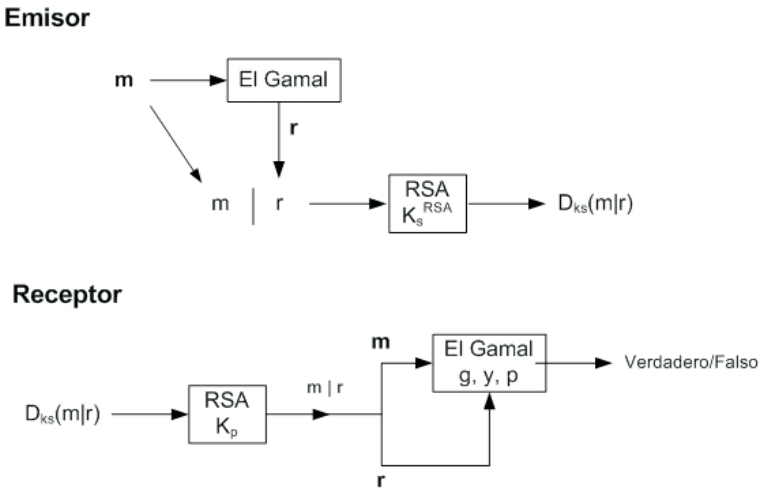


Fig. 3.10: Esquema del sistema de generación/verificación de firma

a) Determinamos la firma pública (g, y, p). y=$15^2 \mod 23 = 18$

- Mensaje $m = 6$.
- Se obtiene un aleatorio $z = 3$ coprimo con $p - 1 = 22$.
- Se deriva un valor: $a = g^z \mod p$
- Se halla $b \in z_{p-1}$, que verifica: $m = (x \cdot a + z \cdot b) \mod p - 1$

b) Calculamos a: $a = 15^3 \mod 23 = 17$

c) Hallamos b' tal que $1 = z \cdot b' \mod p - 1$. De forma equivalente:

$$1 = z \cdot b' + k \cdot (p - 1)$$

$$1 = 3 \cdot b' + 22 \cdot k \implies \begin{cases} k = 1 \\ b' = -7 \equiv 15 \mod 22 \end{cases}$$

d) Dado que $b = (m - x \cdot a) \cdot b' \mod p - 1$, entonces: $b = (6 - 2 \cdot 17) \cdot 15 \mod 22 = 20$

e) $r = a|b = 17|20$

Comprobación, si se verifica $y^a \, a^b = g^m \mod p$, el resumen es correcto.

f) Comprobación:

$$y^a\, a^b = g^m \mod p$$
$$18^{17} \cdot 17^{20} = 15^6 \mod 23$$
$$18^{17} = 18^{10001_2} = ((18^2)^2)^2 \cdot 18 \equiv 8 \mod 23$$
$$17^{20} = 17^{10100_2} = (((17^2)^2 \cdot 17)^2)^2 \equiv 16 \mod 23$$

Se verifica:

$$\left.\begin{array}{l} y^a \cdot a^b = 8 \cdot 16 \mod 23 = 13 \\ g^m \equiv 15^6 \mod 23 = 13 \end{array}\right\} OK$$

g) Validez de la función resumen propuesta:

- El resumen es de longitud fija, con un valor de bits necesario para concatenar $a|b$.
- Dado m, es fácil calcular r, aunque la exponenciación empleada puede ser computacionalmente lenta en algunos casos.
- Dado r, es imposible en la práctica, hallar m si no se conoce x.
- Es poco probable que dos mensajes, m y m', den lugar al mismo r. Se puede controlar la probabilidad en función del tamaño de m máximo y del valor de p.
- Dado un m, es prácticamente imposible hallar otro m' que cumpla $r(m) = r(m')$ si no se conoce x.

Problema 10

Se quiere llevar a cabo un cifrador bloque de cuatro bits mediante un LFSR de longitud 4. Para ello, se carga como estado del LFSR el cuarteto de los bits a cifrar y se hace evolucionar **k** ciclos, y el estado resultante es el valor del cuarteto cifrado. Como polinomio de conexiones, se utiliza un valor **C(D)** fijo para todos los valores de **k**.

Se pide:

a) Si **C(D)** es primitivo ¿cuál es el número de claves distintas?

b) Para **k**=4, el cifrado de [0 0 0 1] (en todas las ternas, el mayor peso se halla a la izquierda) es [0 0 1 1]. ¿Cuánto vale **C(D)**?

c) Para **k**=7, ¿cuánto vale el cifrado de [0 0 0 1]?

d) Para **k**=7, el cifrado del mensaje [0 0 0 1] [0 0 0 1] [0 0 0 1] es [1 0 1 1] [0 0 1 0] [1 1 1 0]. Razone por qué puede asegurarse que el cifrado no se está usando en modo nativo o ECB.

e) Sabiendo que se trata de un cifrado CBC, ¿cuál es el vector inicial?

Solución

a) Visto en forma polinómica, $\text{Cif}(M) = D^k M(D) \mod C(D)$. Si C(D) es primitivo, se alcanzan todos los estados menos el 0; por tanto, existen $2^4 - 1 = 15$ claves distintas.

b) Del enunciado, se deduce que para $k = 4, D^4 \cdot 1 \mod C(D) = D+1$. Supóngase que $C(D) = D^4 + aD^3 + bD^2 + cD + 1$, donde a, b y c son desconocidos. De la ecuación anterior, pueden encontrarse planteando la división:

$$\begin{array}{l|l} D^4 & D^4+aD^3+bD^2+cD+1 \\ \hline D^4+aD^3+bD^2+cD+1 & 1 \\ \hline aD^3+bD^2+cD+1 \end{array} = D+1 \rightarrow a=0\ ;\ b=0\ ;\ c=1 \rightarrow C(D)= D^4+D+1$$

de donde se deduce que $C(D) = D^4 + D + 1$.

c) Del enunciado, se deduce que para $k = 7$ se ha de realizar la operación: $D^7 \cdot 1 = \mod D^4 + D + 1$. Por tanto, se tiene:

$$D^4 \cdot D^3 = (D+1)D^3 = D^4 + D^3 = D^3 + D + 1 = [1\ 0\ 1\ 1]$$

d) Porque en ECB los bloques iguales provocan cifrados iguales y, en este caso, no es así.

e) Del apartado *c*, se tiene que para $k = 7$ el cifrado de [0 0 0 1] es [1 0 1 1], que constituye el primer bloque de texto claro y texto cifrado del encadenamiento CBC. De esto se deduce que el vector inicial ha de ser [0 0 0 0].

Problema 11

En un sistema de comunicaciones inalámbrico, los terminales se autentican utilizando un servidor central. Los terminales intercambian entre sí mensajes cortos de forma confidencial, utilizando el algoritmo RSA. Dado que los terminales no disponen de claves públicas, se genera de forma dinámica para cada sesión, entre terminales A y B, una clave $K_{P_{AB}} = (e_{AB}, n_{AB})$, donde e_{AB} es de valor constante 11 y n_{AB} es el producto de dos primos, p_{AB} y q_{AB}. Los valores de dichos números primos se derivan utilizando, para cada uno de ellos, el mecanismo de operación Diffie-Hellman para compartir un secreto.

El intercambio de mensajes entre los terminales A y B para la compartición de un secreto (p_{AB} y q_{AB}) se realiza utilizando el servidor central como intermediario. De esta forma, se garantiza la identidad entre los terminales. Los mensajes intercambiados se

envían desde el terminal al servidor de forma confidencial utilizando la clave pública del servidor, K_{Serv}, correspondiente al algoritmo RSA. Cuando el servidor recibe el criptograma enviado por un terminal, lo descifra y lo retransmite al otro terminal.

Considerando que:

1. La operación del mecanismo Diffie-Hellman utiliza: $a = 5$ y $p = 97$.
2. Los valores aleatorios generados por los terminales para el secreto compartido de cada número primo son:
   - p_{AB}: terminal A genera $x_1 = 2$; terminal B genera $y_1 = 5$
   - q_{AB}: terminal A genera $x_2 = 7$; terminal B genera $y_2 = 10$
3. La clave pública del servidor es $K_{\text{Serv}} = (e, n) = (3, 319)$.

Determine:

a) El valor de los mensajes cifrados con RSA que se envían desde el terminal A al servidor para la generación de p_{AB} y q_{AB}, respectivamente.

b) Los mensajes enviados en claro desde el servidor al terminal A para la generación de p_{AB} y q_{AB}, respectivamente.

c) La clave pública $K_{P_{AB}}$ de la sesión RSA entre los terminales A y B a partir de los mensajes recibidos por el terminal A y los números aleatorios generados por dicho terminal.

d) La clave secreta de la sesión RSA entre los terminales.

e) El criptograma enviado del terminal A al B cuando el mensaje en claro es 9.

Solución

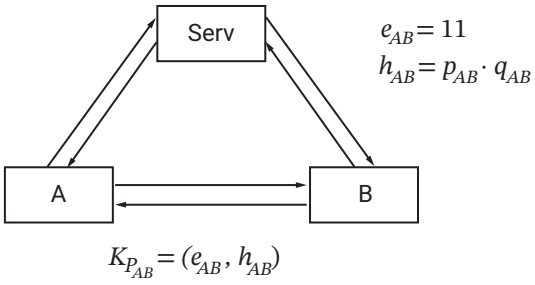


Fig. 3.11: Esquema propuesto

1. Diffie-Hellman $q = 5$ y $p = 97$

2.

	p_{AB}	q_{AB}
A	$x_1 = 2$	$x_2 = 7$
B	$y_1 = 5$	$y_2 = 10$

3. $K_{\text{serv}} = (e, n) = (3, 319)$

a) Mensajes cifrados RSA de A al servidor. Para p_{AB}:

$m_1^A = a^{x_1} \mod p = 5^2 \mod 97 = 25$

Para q_{AB}:

$m_2^A = a^{x_2} \mod p = 5^7 \mod 97 = 40$

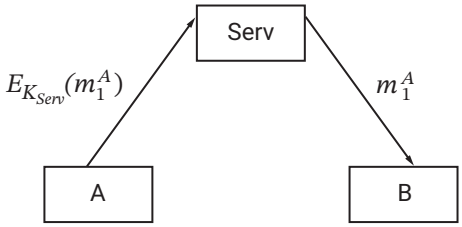


Fig. 3.12: Mensajes RSA cifrados de A al servidor

Criptogramas:

$c_1^A = (m_1^A)^e \mod n = 25^3 \mod 319 = 313$

$c_2^A = (m_2^A)^e \mod n = 40^3 \mod 319 = 200$

b) Para p_{AB}.

$m_1^B = a^{y_1} \mod p = 5^5 \mod 97 = 21$

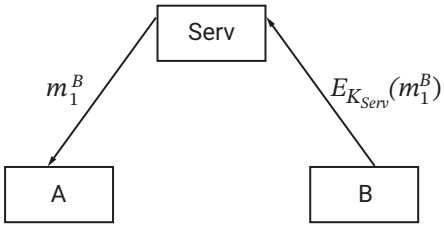


Fig. 3.13: Mensajes enviados en claro del servidor al terminal A

Para q_{AB}.

$m_2^B = a^{y_2} \mod p = 5^10 \mod 97 = 53$

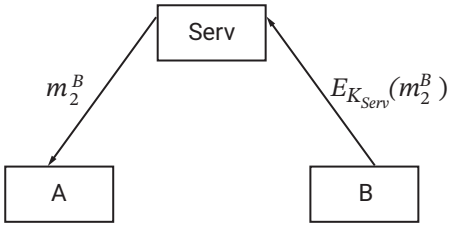


Fig. 3.14: Mensajes enviados en claro del servidor al terminal A

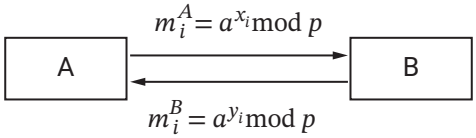


Fig. 3.15: Intercambio Diffie-Hllman de un secreto

c) Indirectamente, entre A y B se ha hecho el intercambio.

El secreto compartido es:

$s_i = (m_i^B)^{y_i} \mod p = (m_i^B)^{x_i} \mod p = a^{x_i y_i} \mod p$

$s_1 = p_{AB} = (m_1^B)^{y_1} \mod p = 21^2 \mod 97 = 53$

$s_2 = q_{AB} = (m_2^B)^{y_2} \mod p = 53^7 \mod 97 = 3$

$n_{AB} = p_{AB} \cdot q_{AB} = 53 \cdot 3 = 159$

$K_{p_{AB}} = (e_{AB}, n_{AB}) = (11, 159)$

d) $K_s = (d, n)$

$e \cdot d + k\Phi(n) = 1$

$\Phi(n) = (p_{AB} - 1)(q_{AB} - 1) = 52 \cdot 2 = 104$

$11d + 104k = 1$

Algoritmo de Euclides extendido:

1. $104 \cdot 1 + 11 \cdot 0 = 104 \qquad (104 = 11 \cdot 9 + 5)$
2. $104 \cdot 0 + 11 \cdot 1 = 11 \qquad (11 = 5 \cdot 2 + 1)$
3. $104 + (-9) \cdot 11 = 5$
4. $(-2) \cdot 104 + (1 + 18) \cdot 11 = 1$
   $k = -2$
   $d = 10 \mod 104 = 19 \quad \Rightarrow \quad K_{s_{AB}} = (19, 159)$

e) $c = m^{e_{AB}} \mod n_{AB} = 9^{11} \mod 159 = 123$

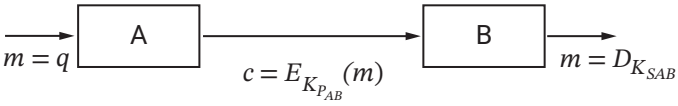


Fig. 3.16: Criptograma enviado

Problema 12

Un cifrador en flujo consta de un LFSR y una función de salida. Trabajando en modo síncrono, se observa que la secuencia generada tiene un período de 2.047 bits.

a) Si el diseñador indica que el número de celdas es < 11, ¿podemos estar seguros de que miente?

b) ¿Podríamos afirmar que el número de celdas es 11? Posteriormente, se configura otro cifrador (también basado en LFSR y con una función de salida) para que opere en modo autosincronizante. Cuando se han obtenido 2.047 bits, no se ha encontrado ningún período.

c) ¿Es posible que el número de celdas del LFSR sea < 11?

Solución

a) Si el número de celdas es < 11, el período sería, como máximo, $2^{10} - 1 = 1.023$, ya que la función de salida no aumenta dicho período. Por tanto, podemos estar seguros de que, con un número de celdas ≤ 10, no se puede conseguir dicho período. Es decir, podemos garantizar que el diseñador miente.

b) No, ya que si el número de celdas es mayor y el polinomio no es primitivo, el período no será $2^L - 1$, siendo L el número de celdas. Es decir, puede ser que el número de celdas sea > 11.

c) Sí, puesto que, cuando el cifrador trabaja en modo autosincronizante, la salida no tiene por qué ser periódica ya que en la realimentación influirán los datos del usuario.

Curricula

Mónica Aguilar Igartua es doctora ingeniera en telecomunicación y profesora titular de universidad en la UPC desde 2001. Pertenece al Grupo de Investigación de Servicios Telemáticos del Departamento de Ingeniería Telemática. Trabaja en servicios para vehículos eléctricos y desarrollo de herramientas para la mejora de la movilidad urbana.

Jordi Forné Muñoz es catedrático de universidad en la UPC desde 2021. Su investigación se centra en el campo de la seguridad y la privacidad de la información.

Jorge Mata Díaz es doctor ingeniero de Telecomunicaciones y profesor titular de universidad en el Departamento de Ingeniería Telemática de la UPC desde 1997. Su investigación se centra en el estudio de algoritmos para la transmisión de información en Internet.

Francisco Rico Novella es doctor ingeniero de Telecomunicación por la UPC (1995). Desde 1997, es profesor titular de universidad adscrito a la Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona (ETSETB). Su investigación se centra en el campo de la criptografía, la codificación y las comunicaciones de corto alcance.

Alfonso Rojas Espinosa es doctor ingeniero de Telecomunicación por la UPC y profesor titular de universidad adscrito a la ETSETB. Sus intereses de investigación están relacionados con la transmisión de datos.

Miquel Soriano Ibáñez es catedrático de universidad en la UPC desde 2007. Su investigación se centra en la seguridad de la información y la protección de los derechos de autor. Ha participado y dirigido numerosos proyectos financiados por la Administración (nacional e internacional) y por empresas privadas.

Transmisión de datos

Problemas resueltos

Los profesores de la asignatura Transmisión de Datos, impartida en la UPC, hemos elaborado este libro de problemas resueltos representativos de la misma. El libro se estructura en tres temas: codificación de fuente, criptografía y codificación de canal.

La transmisión de datos es el conjunto de técnicas y conceptos que surgen al estudiar el problema de la transmisión de información digital, cualquiera que sea su origen, a través de un canal limitado en ancho de banda y potencia. La codificación de fuente contempla la compresión de las fuentes de datos a partir del concepto de información. Los objetivos principales a los que sirve la criptografía son la confidencialidad, la integridad y la autenticidad en el tratamiento de la información en formato electrónico.

Cuando el índice de error del sistema de transmisión sin codificar es demasiado alto, es necesario recurrir a técnicas de codificación de canal, para detectar errores y realizar una retransmisión de los datos, o para corregirlos.

9 788476 535141

iniciativa digital politècnica
Publicacions Acadèmiques de la UPC